

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

QM protein - protein search, using sw model

Run on: September 17, 2004, 11:02:14 : Search time 121 Seconds  
(without alignments)  
200.819 Million cell updates/sec

Title: US-09-625-137-21  
Perfect score: 506  
Sequence: 1 KTCSPKQFACRQITCISKG.....LCNGVQDCWDGSDGPHCRE 86

Scoring table: BLOSUM62

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseq1980s:\*  
2: Geneseq1990s:\*  
3: Geneseq2000s:\*  
4: Geneseq2001s:\*  
5: Geneseq2002s:\*  
6: Geneseq2003as:\*  
7: Geneseq2003bs:\*  
8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	506	100.0	86	5 AAU81033	Aau81033 Human alp
2	506	100.0	790	4 AAM78716	Aam78716 Human pro
3	506	100.0	4544	2 AAR47861	Aar47861 Alpha 2-M
4	506	100.0	4544	2 AAR60517	Aar60517 Human alp
5	506	100.0	4544	4 AAM79091	Aam79091 Human pro
6	506	100.0	4544	5 AAU81019	Aau81019 Human alp
7	506	100.0	4544	6 ABP56839	Abp56839 Human LRP
8	506	100.0	4544	6 ABUS9744	Abu9744 Protein d
9	506	100.0	4544	7 ADD14025	Add14025 Human src
10	506	100.0	4563	4 ABB11353	Abb11353 Human LBL
11	486	96.0	4529	5 AAU81016	Aau81016 Mouse alp
12	486	96.0	4545	5 AAU74797	Aau74797 Mouse alp
13	281	55.3	4636	4 AAE11937	Aae11937 Human Cg1
14	280	55.3	639	4 AAE11928	Aae11928 Human Cg1
15	280	55.3	4599	6 ABP56837	Abp56837 Human LRP
16	275	54.3	165	4 ABG18412	Abg18412 Novel hum
17	258	51.0	44	5 AAU81032	Aau81032 Human alp
18	253	50.0	43	5 AAU81034	Aau81034 Human alp
19	227	44.9	38	6 ABU61280	Abu61280 Human A d
20	224	44.3	37	6 ABU61281	Abu61281 Human A d
21	211.5	41.8	81	5 AAU81064	Aau81064 Human alp
22	211.5	41.8	123	5 AAU81082	Aau81082 Human alp
23	211.5	41.8	170	5 AAU81059	Aau81059 Human alp
24	211.5	41.8	209	5 AAU81056	Aau81056 Human alp
25	211.5	41.8	248	5 AAU81052	Aau81052 Human alp

26	211.5	41.8	289	5 AAU81047	Aau81047 Human alp
27	205.5	40.6	4601	4 ABB59371	Abb59371 Drosophil
28	202	39.9	727	2 AAR05533	Aar05533 Fragment
29	202	39.9	761	4 ABB60973	Abb60973 Drosophil
30	202	39.9	792	4 ABB61029	Abb61029 Drosophil
31	197.5	39.0	1963	4 ABB58053	Abb58053 Drosophil
32	194	38.3	819	6 ADA54122	Ada54122 Human pro
33	193	38.1	904	7 ADD93401	Add93401 Human lip
34	193	38.1	963	6 ABP56838	Abp56838 Human apo
35	193	38.1	1012	5 AAU91285	Aau91285 Human NOV
36	191.5	37.8	159	2 AAR05532	Aar05532 Fragment
37	191.5	37.8	695	5 AAU91286	Aau91286 Human NOV
38	191.5	37.8	699	6 ABUS6579	Abu56579 Lung canc
39	191.5	37.8	729	5 AAU78665	Aau78665 Human NOV
40	191.5	37.8	729	5 AAU91282	Aau91282 Human NOV
41	191.5	37.8	762	5 AAU91283	Aau91283 Human NOV
42	191.5	37.8	762	5 AAU78666	Aau78666 Human NOV
43	191.5	37.8	775	5 ADD93398	Add93398 Human lip
44	191.5	37.8	834	7 ADD93402	Add93402 Human lip
45	191.5	37.8	847	5 AAU91289	Aau91289 Human NOV

ALIGNMENTS

RESULT 1

AAU81033

ID AAU81033 standard; protein; 86 AA.

AC AAU81033;

DT 09-APR-2002 (first entry)

DE Human alpha2 macroglobulin (alpha2M) receptor #2 peptide #2.

KW Human; mouse; alpha2 macroglobulin; receptor; alpha2M; HSP;  
KW heat shock protein; alpha2M receptor-HSP complex; autoimmune disorder;  
KW multiple sclerosis; rheumatoid arthritis; endocytosis; inflammation;  
KW cytokine clearance; antigen presentation disruption; carcinoma; sarcoma;  
KW proliferative disorder; cancer; infectious disease; bacterial infection;  
KW intracellular parasite; hypercholesterolemia; protozoan infection;  
KW Alzheimer's disease; diabetes; osteoporosis; viral infection; protein.

XX Homo sapiens.

XX WO200192474-A1.

XX 06-DEC-2001.

PD 04-JUN-2001; 2001WO-US018041.

PF 02-JUN-2000; 2000US-0209095P.

PR 25-JUL-2000; 2000US-00625137.

PR 22-SEP-2000; 2000US-00668724.

XX 28-DEC-2000; 2000US-00750972.

XX (UYCO-) UNIV CONNECTICUT HEALTH CENT.

PA Srivastava PK;

XX WPI; 2002-122061/16.

PT Screening assays for identifying compounds useful for treating immune

PT disorders, comprises identification of compounds that modulate alpha 2-

XX macroglobulin receptor-heat shock protein interaction.

PS Disclosure; Page 22; 236pp; English.

CC The invention relates to screening assays comprising identification of

CC compounds that modulate alpha2 macroglobulin (alpha2M) receptor (which

CC also functions as a heat shock protein (HSP) receptor)-HSP interaction. A

CC compound that modulates the activity of an alpha2M receptor-HSP complex

CC can be identified by contacting the compound with HSP and alpha2M

CC receptor and measuring the level of alpha2M activity or expression. If  
 CC the level differs from that perceived in the absence of the test  
 CC compound, a compound that modulates an alpha2M receptor-HSP-mediated  
 CC process is identified. The identified compounds are useful for treating  
 CC autoimmune disorders (such as multiple sclerosis or rheumatoid  
 CC arthritis), diseases or disorders involving disruption of antigen  
 CC presentation, endocytosis, cytokine clearance or inflammation,  
 CC proliferative disorders (such as cancers including sarcomas and  
 CC carcinomas), infectious diseases (such as those caused by viruses,  
 CC bacteria, protozoans and intracellular parasites), hypercholesterolaemia,  
 CC Alzheimer's disease, diabetes and osteoporosis. Sequences AAU81016-  
 CC AAU81073 represent human and mouse alpha2M receptors and peptide  
 CC fragments of the invention  
 XX  
 SQ Sequence 86 AA;

Query Match 100.0%; Score 506; DB 5; Length 86;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-43;  
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRDQITCISKGRCDGERDCPDGSDPEAIPCQSKAORCQPNHNCLGTEL 60  
 DB 1 KTCSPKQFACRDQITCISKGRCDGERDCPDGSDPEAIPCQSKAORCQPNHNCLGTEL 60  
 QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86  
 DB 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86

RESULT 2  
 AAM78716  
 ID AAM78716 standard; protein; 790 AA.  
 AC AAM78716;  
 DT 06-NOV-2001 (first entry)  
 XX Human protein SEQ ID NO 1378.  
 DE  
 DE Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorder; arthritis; inflammation.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200157190-A2.  
 XX  
 XX 09-AUG-2001.  
 XX  
 XX 05-FEB-2001; 2001WO-US004098.  
 XX  
 XX 03-FEB-2000; 2000US-00496914.  
 XX 27-APR-2000; 2000US-00560875.  
 XX 20-JUN-2000; 2000US-00598075.  
 XX 19-JUL-2000; 2000US-00620325.  
 XX 01-SEP-2000; 2000US-00654936.  
 XX 15-SEP-2000; 2000US-00663561.  
 XX 20-OCT-2000; 2000US-00693325.  
 XX 30-NOV-2000; 2000US-00728422.  
 XX (HYSB-) HYSQ INC.  
 XX  
 XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;  
 PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
 XX  
 XX WPI; 2001-476283/51.  
 XX N-PSDB; AAK51849.  
 XX  
 XX Nucleic acids encoding polypeptides with cytokine-like activities, useful  
 XX in diagnosis and gene therapy.

PS Claim 20; Page 3637-3639; 6221pp; English.  
 XX  
 XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111  
 CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the  
 CC sequence listing were missing at the time of publication  
 XX  
 SQ Sequence 790 AA;

Query Match 100.0%; Score 506; DB 4; Length 790;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-42;  
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRDQITCISKGRCDGERDCPDGSDPEAIPCQSKAORCQPNHNCLGTEL 60  
 DB 25 KTCSPKQFACRDQITCISKGRCDGERDCPDGSDPEAIPCQSKAORCQPNHNCLGTEL 84  
 QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86  
 DB 85 CVPMSRLCNGVQDCMDGSDGPHCRE 110

RESULT 3  
 AAR47861  
 ID AAR47861 standard; protein; 4544 AA.  
 AC AAR47861;  
 DT 25-MAR-2003 (revised)  
 DT 20-JUL-1994 (first entry)  
 XX  
 XX Alpha 2-Macroglobulin/LDL-receptor related protein.  
 DE  
 XX alpha-2 macroglobulin; Low Density Lipoprotein; LDL receptor family;  
 KW LDL receptor related protein; LRP; small rhinovirus receptor; deriv;  
 KW minor Rhinovirus; alpha2MR/LRP.  
 XX  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH Misc-difference 211..260  
 FT Misc-difference /note= "50 residues not shown in SEQ.ID.No.4"  
 FT Misc-difference /note= "Residue not shown in SEQ.ID.No.4"  
 FT Misc-difference /note= "Residue not shown in SEQ.ID.No.4"  
 FT Misc-difference 3050  
 XX WO9401553-A1.  
 XX  
 XX 20-JAN-1994.  
 XX  
 XX 05-JUL-1993; 93WO-EP001728.  
 XX  
 XX 08-JUL-1992; 92DE-04222385.  
 XX 22-AUG-1992; 92DE-04227892.  
 XX 19-FEB-1993; 93DE-04305063.  
 XX (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
 XX  
 XX Hofer F, Kowalski H, Gruenberger M, Machat H, Huettinger M;  
 PI Kerjaschki D, Kuechler E, Blasas D;  
 XX  
 XX WPI; 1994-035060/04.  
 XX

PT New peptide: derivs. of receptor for rhinovirus - of the small receptor  
PT GP., and derived DNA, transformed cells and antibodies, used e.g. to  
PT treat or prevent rhinovirus infection.  
XX  
XX Claim 5; Fig 2; 76pp; German.

XX  
XX Functional derivatives of members of the Minor Rhinovirus Receptor group  
CC are claimed. The alpha-2 Macroglobulin/LDL-receptor related protein of  
CC sequence AAK47861 (Herz et al. EMBO J. 7;4119-4127 (1988)) is a preferred  
CC parent receptor. The derivs, which are preferably soluble, extracellular  
CC forms of the native receptors, are useful for treating and preventing  
CC viral (esp. rhinoviral) infections. N.B. the SEQ.ID. listing includes a  
CC sequence (no.4) which differs from the alpha2-MR/LRP sequence as  
CC indicated in the Features Table. (Updated on 25-MAR-2003 to correct PN  
CC field.)  
XX

SQ Sequence 4544 AA;  
Query Match 100.0%; Score 506; DB 2; Length 4544;  
Best Local Similarity 100.0%; Pred. No. 1.4e-41;  
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCSPKQFACRDQITCISKGRCDGERDCPDGSDPEAIPCQSKAQRCPNEHNCIGTEL 60  
Db 25 KTCSPKQFACRDQITCISKGRCDGERDCPDGSDPEAIPCQSKAQRCPNEHNCIGTEL 84  
QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86  
Db 85 CVPMSRLCNGVQDCMDGSDGPHCRE 110

RESULT 4  
AAR60517  
ID AAR60517 standard; protein; 4544 AA.  
XX  
XX AAR60517;  
XX  
XX 25-MAR-2003 (revised)  
DT 22-MAR-1995 (first entry)  
XX  
XX Human alpha-2-MR.  
XX  
XX Serine protease; Factor-Xa; recognition site; fusion protein cleavage;  
KW protein folding; alpha-2-MR;  
KW alpha-2-macroglobulin receptor/low density lipoprotein receptor.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO9418227-A2.  
FN  
XX  
XX 18-AUG-1994.  
PD  
XX  
XX 04-FEB-1994; 94WO-DK000054.  
PF  
XX  
XX 04-FEB-1993; 93DK-00000130.  
PR  
XX  
XX 05-FEB-1993; 93DK-00000139.  
PR  
XX  
XX 03-DEC-1993; 93WO-GB002492.  
PR  
XX  
XX (DENZ-) DENZYME APS.  
PA  
XX  
XX Trogersen HC, Holtet TL, Etzerodt M;  
XX  
XX WPI; 1994-279681/34.  
DR  
XX  
XX Refolding of polypeptide molecules - using a cyclic process involving  
XX denaturing and renaturing conditions to produce a correctly folded prod.  
PT  
XX  
XX Disclosure; Page 131-146; 202pp; English.  
PS  
XX  
XX Various domains and domain clusters of human alpha-2-MR protein have been  
CC PCR amplified using the primers given in AAK71252-65. (Updated on 25-MAR-  
CC 2003 to correct PN field.)  
CC

SQ Sequence 4544 AA;  
Query Match 100.0%; Score 506; DB 2; Length 4544;  
Best Local Similarity 100.0%; Pred. No. 1.4e-41;  
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCSPKQFACRDQITCISKGRCDGERDCPDGSDPEAIPCQSKAQRCPNEHNCIGTEL 60  
Db 25 KTCSPKQFACRDQITCISKGRCDGERDCPDGSDPEAIPCQSKAQRCPNEHNCIGTEL 84  
QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86  
Db 85 CVPMSRLCNGVQDCMDGSDGPHCRE 110

RESULT 5  
AAM79091  
ID AAM79091 standard; protein; 4544 AA.  
XX  
XX AAM79091;  
XX  
XX 06-NOV-2001 (first entry)  
DT  
XX  
XX Human protein SEQ ID NO 1753.  
DE  
XX  
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorder; arthritis; inflammation.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200157190-A2.  
FN  
XX  
XX 09-AUG-2001.  
PD  
XX  
XX 05-FEB-2001; 2001WO-US004098.  
PF  
XX  
XX 03-FEB-2000; 2000US-00496914.  
PR  
XX  
XX 27-APR-2000; 2000US-00560875.  
PR  
XX  
XX 20-JUN-2000; 2000US-00598075.  
PR  
XX  
XX 19-JUL-2000; 2000US-00820325.  
PR  
XX  
XX 01-SEP-2000; 2000US-00854936.  
PR  
XX  
XX 15-SEP-2000; 2000US-00663561.  
PR  
XX  
XX 20-OCT-2000; 2000US-00693325.  
PR  
XX  
XX 30-NOV-2000; 2000US-00728422.  
PR  
XX  
XX (HYSE-) HYSEQ INC.  
PA  
XX  
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;  
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
XX  
XX WPI: 2001-476283/51.  
DR  
XX  
XX N-PSDB; AAK52224.  
DR  
XX  
XX Nucleic acids encoding polypeptides with cytokine-like activities, useful  
PT in diagnosis and gene therapy.  
PT  
XX  
XX Claim 20; Page 4086-4095; 6221pp; English.  
PS  
XX  
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
XX encoded polypeptides (AAM78323-AAK80302) that exhibit activity relating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111  
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the

CC sequence listing were missing at the time of publication

XX  
SQ Sequence 4544 AA;  
Query Match 100.0%; Score 506; DB 4; Length 4544;  
Best Local Similarity 100.0%; Pred. No. 1.4e-41;  
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRDQITCISKWRCDEGDCPDGSDPEAIPCQSKAQRCPQNEHNCIGTEL 60  
DB 25 KTCSPKQFACRDQITCISKWRCDEGDCPDGSDPEAIPCQSKAQRCPQNEHNCIGTEL 84  
QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86  
DB 85 CVPMSRLCNGVQDCMDGSDGPHCRE 110

## RESULT 6

AAU81019  
ID AAU81019 standard; protein; 4544 AA.

XX  
AC AAU81019;

DT 09-APR-2002 (first entry)

XX Human alpha2 macroglobulin (alpha2M) receptor #2.

DE Human; mouse; alpha2 macroglobulin; receptor; alpha2M; HSP;  
KW heat shock protein; alpha2M receptor-HSP complex; autoimmune disorder;  
KW multiple sclerosis; rheumatoid arthritis; endocytosis; inflammation;  
KW cytokine clearance; antigen presentation disruption; carcinoma; sarcoma;  
KW proliferative disorder; cancer; infectious disease; bacterial infection;  
KW intracellular parasite; hypercholesterolaemia; protozoan infection;  
KW Alzheimer's disease; diabetes; osteoporosis; viral infection; protein.

XX Homo sapiens.

XX W0200192474-A1.

XX 06-DEC-2001.

XX 04-JUN-2001; 2001WO-US018041.

XX 02-JUN-2000; 2000US-0209095P.

XX 25-JUL-2000; 2000US-00625137.

XX 22-SEP-2000; 2000US-00668724.

XX 28-DEC-2000; 2000US-00750972.

XX (UYCO-) UNIV CONNECTICUT HEALTH CENT.

XX Srivastava PK;

XX WPI; 2002-122061/16.

XX N-PSDB; ABK24096.

XX Screening assays for identifying compounds useful for treating immune disorders, comprises identification of compounds that modulate alpha 2-macroglobulin receptor-heat shock protein interaction.

XX Claim 8; Fig 14B; 236pp; English.

XX The invention relates to screening assays comprising identification of compounds that modulate alpha2 macroglobulin (alpha2M) receptor (which also functions as a heat shock protein (HSP) receptor)-HSP interaction. A compound that modulates the activity of an alpha2M receptor-HSP complex can be identified by contacting the compound with HSP and alpha2M receptor and measuring the level of alpha2M activity or expression. If the level differs from that perceived in the absence of the test compound, a compound that modulates an alpha2M receptor-HSP-mediated process is identified. The identified compounds are useful for treating autoimmune disorders (such as multiple sclerosis or rheumatoid arthritis), diseases or disorders involving disruption of antigen presentation, endocytosis, cytokine clearance or inflammation,

CC proliferative disorders (such as cancers including sarcomas and carcinomas), infectious diseases (such as those caused by viruses, bacteria, protozoans and intracellular parasites), hypercholesterolaemia, Alzheimer's disease, diabetes and osteoporosis. Sequences AAU81016-AAU81073 represent human and mouse alpha2M receptors and peptide fragments of the invention

XX Sequence 4544 AA;

Query Match 100.0%; Score 506; DB 5; Length 4544;  
Best Local Similarity 100.0%; Pred. No. 1.4e-41;  
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRDQITCISKWRCDEGDCPDGSDPEAIPCQSKAQRCPQNEHNCIGTEL 60  
DB 25 KTCSPKQFACRDQITCISKWRCDEGDCPDGSDPEAIPCQSKAQRCPQNEHNCIGTEL 84

QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86

DB 85 CVPMSRLCNGVQDCMDGSDGPHCRE 110

## RESULT 7

ABP56839

ID ABP56839 standard; protein; 4544 AA.

XX  
AC ABP56839;

DT 03-APR-2003 (first entry)

XX Human LRP protein SEQ ID NO:6.

DE Human; sex hormone binding globin; SHBG; MEGF-7; megalin; LRP1B;  
KW low density lipoprotein receptor related protein; LDL; VLDL receptor;  
KW very low density lipoprotein receptor; apolipoprotein E receptor 2;  
KW cubulin; steroid hormone; steroid hormone binding protein; cytostatic;  
KW contraceptive; androgen; oestrogen; progesterone; corticoid; testosterone;  
KW dihydrotestosterone; oestradiol; prostate cancer; breast cancer.

XX Homo sapiens.

XX W0200299438-A2.

XX 12-DEC-2002.

XX 04-JUN-2002; 2002WO-DK000379.

XX 07-JUN-2001; 2001DK-00000887.

XX 08-JUN-2001; 2001US-0296489P.

XX (DELB-) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.

XX (UYAA-) UNIV AARHUS.

XX Willnow T, Nykjaer A;

XX WPI; 2003-140657/13.

XX Screening for a compound that alters uptake of steroid hormone (SH) into cells presenting SH binding protein receptor, by identifying compound that alters uptake by the receptor of SH bound to SH binding protein.

XX Claim 21; Page 151-171; 188pp; English.

XX The present invention describes a method (M1) of screening for a compound which alters uptake of steroid hormones (SH) into cells presenting a SH binding protein (SBP) receptor (I). The method involves providing an assay for measuring binding to or uptake by (I) of SH bound to or complexed with SBP, adding the compound to be tested to the assay, and determining the amount of SH bound to or complexed with SBP which is bound to or taken up by (I), comparing the amount determined with an amount measured in the absence of the compound to be tested, where a difference in the two amounts identifies a compound which alters the binding or uptake of SH bound to or complexed with SBP. Also described is

CC a compound (II) which alters the uptake of SH into cells presenting (I) as identified by (M1). (I) has cytostatic and contraceptive activities.  
CC (M1) is useful for screening for a compound which alters uptake of SH.  
CC such as androgens, oestrogens, progestogens and corticoids, preferably  
CC androgens and oestrogens (e.g. testosterone, dihydrotestosterone, and  
CC oestradiol), into cells presenting (I). (II) is useful for treating a  
CC clinical condition such as prostate or breast cancer in an individual.  
CC (II) is also useful for preventing pregnancy in a male. (II) is useful  
CC for preparing a pharmaceutical composition for treating a clinical  
CC condition associated with a steroid hormone, e.g. prostate cancer or  
CC breast cancer. The present sequence represents the specifically claimed  
CC human low density lipoprotein receptor related protein (LRP) protein,  
CC which is given in the exemplification of the present invention  
XX  
SQ Sequence 4544 AA;  
  
Query Match 100.0%; Score 506; DB 6; Length 4544;  
Best Local Similarity 100.0%; Pred. No. 1.4e-41;  
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 KTCSPKQFACRQDQITCISKWRCGERDCPDGSDPEAIPCQSKAQCQPNEHNCIGTEL 60  
25 KTCSPKQFACRQDQITCISKWRCGERDCPDGSDPEAIPCQSKAQCQPNEHNCIGTEL 84  
Db  
QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86  
85 CVPMSRLCNGVQDCMDGSDGPHCRE 110  
Db  
  
RESULT 8  
ABU89744  
ID ABU89744 standard; protein; 4544 AA.  
AC ABU89744;  
XX  
XX 10-JUL-2003 (first entry)  
XX  
XX Protein differentially expressed in cardiovascular disease #38.  
XX  
XX Cardiovascular disease; arteriosclerosis; ischaemia; angina pectoris;  
KW myocardial infarction; cardiant; antiarteriosclerotic; antianginal;  
KW gene therapy; differential gene expression.  
XX  
XX Homo sapiens.  
XX  
XX WO2003031650-A2.  
PN  
XX 17-APR-2003.  
PD  
XX 02-OCT-2002; 2002WO-EP011034.  
PF  
XX 08-OCT-2001; 2001GB-00024145.  
PR  
XX (FARB ) BAYER AG.  
PA  
XX Munnes M, Gehrman M, Wick M, Schmitz G;  
PI  
XX WPI; 2003-403108/38.  
DR  
XX N-PSDB; ACA89917.  
DR  
XX Predicting, diagnosing or prognosing a cardiovascular disease, e.g.  
PT angina, ischemia, myocardial infarction or arteriosclerosis by detection  
PT of a polynucleotide in a biological sample comprises detecting a  
PT hybridization complex.  
XX  
XX Claim 3; Page 353-373; 454pp; English.  
PS  
XX The invention describes a method of predicting, diagnosing or prognosing  
CC a cardiovascular disease by detection of a polynucleotide in a biological  
CC sample comprises hybridising at least one of the polynucleotide to a  
CC nucleic acid material of a biological sample, thus forming a  
CC hybridisation complex, and detecting the hybridisation complex. The  
CC polynucleotides, polypeptides, antisense molecule, antibody and reagent

CC are useful for preparing compositions for preventing, predicting or  
CC diagnosing, or a medicament for treating a cardiovascular disease, e.g.  
CC arteriosclerosis, ischaemia, angina pectoris, or myocardial infarction.  
CC This sequence represents a protein identified in the invention a being  
CC differentially expressed in individuals with cardiovascular disease  
XX  
SQ Sequence 4544 AA;  
  
Query Match 100.0%; Score 506; DB 6; Length 4544;  
Best Local Similarity 100.0%; Pred. No. 1.4e-41;  
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 KTCSPKQFACRQDQITCISKWRCGERDCPDGSDPEAIPCQSKAQCQPNEHNCIGTEL 60  
25 KTCSPKQFACRQDQITCISKWRCGERDCPDGSDPEAIPCQSKAQCQPNEHNCIGTEL 84  
Db  
QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86  
85 CVPMSRLCNGVQDCMDGSDGPHCRE 110  
Db  
  
RESULT 9  
ADD14025  
ID ADD14025 standard; protein; 4544 AA.  
XX  
XX ADD14025;  
AC  
XX 01-JAN-2004 (first entry)  
DT  
XX Human src biomarker polypeptide SEQ ID NO:214.  
DE  
XX predictor set; protein tyrosine kinase activity modulator;  
KW protein tyrosine kinase pathway; protein tyrosine kinase; cytostatic;  
KW gene therapy; drug sensitivity; genetic profile; cancer; human.  
XX  
XX Homo sapiens.  
OS  
XX WO2003062395-A2.  
PN  
XX 31-JUL-2003.  
PD  
XX 17-JAN-2003; 2003WO-US001981.  
PF  
XX 18-JAN-2002; 2002US-0350061P.  
PR  
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
PA  
XX Huang F, Fairchild CR, Lee FY, Shaw P;  
PI  
XX WPI; 2003-636735/60.  
DR  
XX N-PSDB; ADD14619.  
DR  
XX New polynucleotides and polypeptides for predicting the activity of  
PT compounds that interact with protein tyrosine kinases and/or protein  
PT tyrosine kinase pathways.  
PT  
XX Claim 10; SEQ ID NO 214; 139pp; English.  
PS  
XX The present invention describes a predictor set comprising a plurality of  
CC polynucleotides or polypeptides whose expression pattern is predictive of  
CC the response of cells to treatment with a compound that modulates protein  
CC tyrosine kinase activity or members of the protein tyrosine kinase  
CC pathway. Also described: (1) predicting whether a compound is capable of  
CC modulating the activity of cells, comprising obtaining a sample of cells,  
CC determining whether the cells express a plurality of markers, and  
CC correlating the expression of the markers to the compound's ability to  
CC modulate the activity of the cells; (2) a plurality of cell lines for  
CC identifying polynucleotides and polypeptides whose expression levels  
CC correlate with compound sensitivity or resistance of cells associated  
CC with a disease state; and (3) identifying polynucleotides and  
CC polypeptides that predict compound sensitivity or resistance of cells  
CC associated with a disease state, comprising subjecting the plurality of  
CC cell lines to one or more compounds, analysing the expression pattern of

CC a microarray of polynucleotides or polypeptides, and selecting  
CC polynucleotides or polypeptides that predict the sensitivity or  
CC resistance of cells associated with a disease state by using the  
CC expression pattern of the microarray. The polynucleotides and  
CC polypeptides have cytostatic activities, and can be used in gene therapy.  
CC The polynucleotides and polypeptides are useful in predicting the  
CC activity of compounds that interact with protein tyrosine kinases and/or  
CC protein tyrosine kinase pathways. These may be used in determining drug  
CC sensitivity in patients to allow the development of individualized  
CC genetic profiles which aid in treating diseases and disorders (e.g.  
CC cancer) based on patient response at a molecular level. The present  
CC sequence is used in the exemplification of the present invention.

XX SQ Sequence 4544 AA;  
Query Match 100.0%; Score 506; DB 7; Length 4544;  
Best Local Similarity 100.0%; Pred. No. 1.4e-41;  
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRDQITCISKWRCDEGDCPDGSDDEAPEICQSKAQRCQPNHNCLGTEL 60  
DB 25 KTCSPKQFACRDQITCISKWRCDEGDCPDGSDDEAPEICQSKAQRCQPNHNCLGTEL 84  
QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86  
DB 85 CVPMSRLCNGVQDCMDGSDGPHCRE 110

RESULT 10  
ID ABB11353  
XX AC ABB11353 standard; peptide; 4563 AA.  
XX AC ABB11353;  
XX DT 11-JAN-2002 (first entry)  
XX DE Human LDL receptor precursor homologue, SEQ ID NO:1723.  
XX KW Human; cytokine; cell proliferation; tissue growth; immunomodulator; growth factor;  
KW hematopoiesis regulation; tissue growth; immunomodulator; actin;  
KW limbin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;  
KW proliferation; metastasis; cancer; tumour; hematopoietic disorder;  
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
KW chronic inflammatory condition; proliferative retinopathy;  
KW atherosclerosis; coronary heart disease; arterial ischaemia;  
KW bone disorder; osteoporosis; vascular growth disorder;  
KW tissue regeneration; wound healing; infection; immune disorder;  
KW cell culture; drug screening; gene therapy; anti-inflammatory;  
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;  
KW cytoskeletal; osteopathic; vasotropic; cardiac; virucide; antibacterial;  
KW antifungal; vulnery; antiulcer.

OS Homo sapiens.  
XX WO200157188-A2.  
XX PD 09-AUG-2001.  
XX PF 05-FEB-2001; 2001WO-US003800.  
XX PR 03-FEB-2000; 2000US-00496914.  
XX FR 27-APR-2000; 2000US-00560875.  
XX PA (HYSE-) HYSEQ INC.  
XX PI Tang YT, Liu C, Drmanac RT;  
XX DR WPI; 2001-457740/49.  
XX DR N-PSDB; ABA08597.  
XX PT Human proteins and DNA encoding sequences useful for preventing, treating  
PT or ameliorating a medical condition in a mammalian subject e.g. arthritis  
PT and cancer.

XX PS Claim 20; Page 169-172; 1963pp; English.  
CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and  
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The  
CC invention also relates to vectors and recombinant host cells comprising a  
CC nucleotide of the invention, methods of producing the novel polypeptides,  
CC antibodies against the polypeptides, methods of detecting the nucleotides  
CC or polypeptides in a sample, and methods of identifying compounds which  
CC bind to polypeptides of the invention. Although novel, many of the  
CC polypeptides of the invention have homology to known proteins, thereby  
CC giving an insight into their probable biological activities, and hence  
CC potential therapeutic applications. The polypeptides of the invention may  
CC have various activities, including cytokine, cell proliferation or cell  
CC differentiation activities; stem cell growth factor activity;  
CC immunomodulatory activity; tissue growth activity;  
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or  
CC thrombolytic activities; receptor or ligand activities; or may be  
CC involved in oncogenesis, cancer cell proliferation or metastasis.  
CC Depending on their biological activities, polypeptides and nucleotides of  
CC the invention are useful for preventing, treating or ameliorating medical  
CC conditions, e.g., by protein or gene therapy. Such conditions include  
CC cancers, hematopoietic disorders (e.g., myeloid or lymphoid cell  
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),  
CC proliferative retinopathy, atherosclerosis, coronary heart disease,  
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal  
CC vascular growth. Polypeptides involved with tissue regeneration and  
CC repair (or nucleic acids encoding them) may be used to promote wound  
CC healing (e.g., of burns, incisions and ulcers), while those with  
CC immunomodulatory activities may be used in the treatment of viral,  
CC bacterial and fungal infections in addition to immune disorders.  
CC Polypeptides with growth factor activity may be used in cell cultures to  
CC promote cell growth. For example, such polypeptides may be used to  
CC manipulate stem cells in culture to give rise to neuroepithelial cells  
CC that can be used to augment or replace cells damaged by illness,  
CC autoimmune disease or accidental damage. The polypeptides and nucleotides  
CC may also be used in the diagnosis of the above conditions, and in drug  
CC screening techniques. The present sequence represents a novel human  
CC polypeptide of the invention

XX SQ Sequence 4563 AA;  
Query Match 100.0%; Score 506; DB 4; Length 4563;  
Best Local Similarity 100.0%; Pred. No. 1.4e-41;  
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRDQITCISKWRCDEGDCPDGSDDEAPEICQSKAQRCQPNHNCLGTEL 60  
DB 44 KTCSPKQFACRDQITCISKWRCDEGDCPDGSDDEAPEICQSKAQRCQPNHNCLGTEL 103  
QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86  
DB 104 CVPMSRLCNGVQDCMDGSDGPHCRE 129

RESULT 11  
ID AAU81016 standard; protein; 4529 AA.  
XX AC AAU81016;  
XX DT 09-APR-2002 (first entry)  
XX DE Mouse alpha2 macroglobulin (alpha2M) receptor.  
XX KW Human; mouse; alpha2 macroglobulin; receptor; alpha2M; HSP;  
KW heat shock protein; alpha2M receptor-HSP complex; autoimmune disorder;  
KW multiple sclerosis; rheumatoid arthritis; endocytosis; inflammation;  
KW cytokine clearance; antigen presentation disruption; carcinoma; sarcoma;  
KW proliferative disorder; cancer; infectious disease; bacterial infection;  
KW intracellular parasite; hypercholesterolemia; protozoan infection;  
KW Alzheimer's disease; diabetes; osteoporosis; viral infection; protein.

```

XX OS Mus musculus.
XX PN WO2001192474-A1.
XX PD 06-DEC-2001.
XX PF 04-JUN-2001; 2001WO-US018041.
XX PR 02-JUN-2000; 2000US-0209095P.
XX PR 25-JUL-2000; 2000US-00625137.
XX PR 22-SEP-2000; 2000US-00668724.
XX PR 28-DEC-2000; 2000US-00750972.
XX PA (UYCO-) UNIV CONNECTICUT HEALTH CENT.
XX PI Srivastava PK;
XX DR WPI; 2002-122061/16.
XX DR N-PSDB; ABK24094.
XX PT Screening assays for identifying compounds useful for treating immune
XX PT disorders, comprises identification of compounds that modulate alpha 2-
XX PT macroglobulin receptor-heat shock protein interaction.
XX PS Disclosure; Fig 12B; 236pp; English.
XX CC The invention relates to screening assays comprising identification of
XX CC compounds that modulate alpha2 macroglobulin (alpha2M) receptor (which
XX CC also functions as a heat shock protein (HSP) receptor)-HSP interaction. A
XX CC compound that modulates the activity of an alpha2M receptor-HSP complex
XX CC can be identified by contacting the compound with HSP and alpha2M
XX CC receptor and measuring the level of alpha2M activity or expression. If
XX CC the level differs from that perceived in the absence of the test
XX CC compound, a compound that modulates an alpha2M receptor-HSP-mediated
XX CC process is identified. The identified compounds are useful for treating
XX CC autoimmune disorders (such as multiple sclerosis or rheumatoid
XX CC arthritis), diseases or disorders involving disruption of antigen
XX CC presentation, endocytosis, cytokine clearance or inflammation,
XX CC proliferative disorders (such as cancers including sarcomas and
XX CC carcinomas), infectious diseases (such as those caused by viruses,
XX CC bacteria, protozoans and intracellular parasites), hypercholesterolaemia,
XX CC Alzheimer's disease, diabetes and osteoporosis. Sequences AAU81016-
XX CC AAU81073 represent human and mouse alpha2M receptors and peptide
XX CC fragments of the invention
XX SQ Sequence 4529 AA;
Query Match 96.0%; Score 486; DB 5; Length 4529;
Best Local Similarity 95.3%; Pred. No. 1.4e-39;
Matches 82; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 KTCSPKQFACRDQITCISKGRCDGERDCPDGSDPEAIPCQSKAQRCPNEHNCIGTEL 60
Db 26 KTCSPKQFACRDQITCISKGRCDGERDCPDGSDPEAIPCQSKAQRCPNEHNCIGTEL 85
QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86
Db 86 CVPMSRLCNGIQDCMDGSDGEGHCRE 111
RESULT 12
AAU74797
ID AAU74797 standard; protein; 4545 AA.
XX AC AAU74797;
XX DT 09-APR-2002 (first entry)
XX DE Mouse alpha 2 macroglobulin (alpha2MR).
XX KW Alpha 2 macroglobulin; alpha2MR; mouse; immunosuppressive; cytostatic;
KW virucide; antibacterial; protozoacide; antiparasitic; vaccine;
immunotherapy; autoimmune disease; infectious disease;
cell proliferation disorder; systemic lupus erythematosus; IDDM;
insulin dependent diabetes mellitus; multiple sclerosis; scleroderma;
Sjogren's syndrome; myasthenia gravis; chronic active hepatitis;
ulcerative colitis; sarcoma; fibrosarcoma; myxosarcoma; synovium;
carcinoma; hepatoma; Wilm's tumour; cervical cancer; leukaemia;
acute lymphocytic leukaemia; chronic myelocytic leukaemia;
polycythaemia vera; lymphoma; Waldenstrom's macroglobulinaemia;
Hodgkin's disease.
XX OS Mus musculus.
XX PN WO200191787-A1.
XX PD 06-DEC-2001.
XX PF 04-JUN-2001; 2001WO-US018047.
XX PR 02-JUN-2000; 2000US-0209266P.
XX PR 25-JUL-2000; 2000US-00625139.
XX PA (UYCO-) UNIV CONNECTICUT HEALTH CENT.
XX PI Srivastava PK;
XX DR WPI; 2002-114309/15.
XX DR N-PSDB; ABK12951.
XX PT New compositions comprising an alpha (2) macroglobulin-antigenic molecule
XX PT complex, useful for diagnosing, treating or preventing autoimmune
XX PT diseases, proliferative disorders and infectious diseases.
XX PS Disclosure; Fig 6A; 160pp; English.
XX CC The invention describes new pharmaceutical compositions comprising a
XX CC molecular complex or a fusion protein for treating or preventing an
XX CC infectious disease or cancer, and a carrier. The alpha (2) macroglobulin
XX CC antigenic complexes and compositions comprising the polypeptide are
XX CC useful for diagnosing, treating or preventing autoimmune diseases
XX CC (including multiple sclerosis, systemic lupus erythematosus, Sjogren's
XX CC syndrome, insulin dependent diabetes mellitus (IDDM), myasthenia gravis,
XX CC scleroderma, chronic active hepatitis, ulcerative colitis), proliferative
XX CC disorders (including sarcoma e.g. fibrosarcoma, myxosarcoma and synovium
XX CC ; carcinomas e.g. hepatoma, Wilm's tumour and cervical cancer; leukaemias
XX CC e.g. acute lymphocytic leukaemia, chronic myelocytic leukaemia and
XX CC polycythaemia vera; lymphoma e.g. Hodgkin's disease and Waldenstrom's
XX CC macroglobulinaemia), and infectious diseases such as those caused by
XX CC viruses, bacteria, protozoans and intracellular parasites. This is the
XX CC amino acid sequence of the mouse alpha 2 macroglobulin (alpha2MR) used
XX CC for creating the complexes for immunotherapy described in the method of
XX CC the invention
XX SQ Sequence 4545 AA;
Query Match 96.0%; Score 486; DB 5; Length 4545;
Best Local Similarity 95.3%; Pred. No. 1.4e-39;
Matches 82; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 KTCSPKQFACRDQITCISKGRCDGERDCPDGSDPEAIPCQSKAQRCPNEHNCIGTEL 60
Db 26 KTCSPKQFACRDQITCISKGRCDGERDCPDGSDPEAIPCQSKAQRCPNEHNCIGTEL 85
QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86
Db 86 CVPMSRLCNGIQDCMDGSDGEGHCRE 111
RESULT 13
AAE11937
ID AAE11937 standard; protein; 4636 AA.
XX AC AAE11937;
XX DT
XX DE
XX KW
KW

```





Query Match 55.3%; Score 280; DB 4; Length 639;  
 Best Local Similarity 53.6%; Pred. No. 9.2e-20;  
 Matches 45; Conservative 12; Mismatches 27; Indels 0; Gaps 0;

QY 3 CSPKQACRQITICISKGWCDGRCPCDGSDEAPEICQSKAQRCPNEHNCIGTGLCV 62  
 Db 69 CDGFEFLCHDHVTCVQSQWLCDGDPDSDSDSLDTCPEEVEIKPLNHIACLGTKVCV 128  
 QY 63 PMSRLNGVQDCMDGSDGPHCRE 86  
 Db 129 HLSQLCNGVLDCPDGYDEGVHCQE 152

RESULT 15  
 ABP56837

ID ABP56837 standard; protein; 4599 AA.

AC ABP56837;

DT 03-APR-2003 (first entry)

DE Human LRP1B protein SEQ ID NO:4.

XX Human; sex hormone binding globin; SHBG; MEGF-7; megalin; LRP1B;  
 KW low density lipoprotein receptor related protein; LDL; VLDL receptor;  
 KW very low density lipoprotein receptor; apolipoprotein E receptor 2;  
 KW cubulin; steroid hormone; steroid hormone binding protein; cytosolic;  
 KW contraceptive; androgen; oestrogen; progesterone; corticoid; testosterone;  
 KW dihydrotestosterone; oestradiol; prostate cancer; breast cancer.  
 XX Homo sapiens.

XX Key Location/Qualifiers

FH Misc-difference 4343

FT /note= "any amino acid"

XX WO200299438-A2.

XX 12-DEC-2002.

XX 04-JUN-2002; 2002WO-DK000379.

XX 07-JUN-2001; 2001DK-00000887.

XX 08-JUN-2001; 2001US-0296489P.

XX (DELB-) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.

XX (UYAA-) UNIV ARHUS.

XX Willnow T, Nykjaer A;

XX WPI; 2003-140657/13.

XX Screening for a compound that alters uptake of steroid hormone (SH) into  
 PT cells presenting SH binding protein receptor, by identifying compound  
 PT that alters uptake by the receptor of SH bound to SH binding protein.

XX Claim 21; Page 126-146; 188pp; English.

XX The present invention describes a method (M1) of screening for a compound  
 CC which alters uptake of steroid hormones (SH) into cells presenting a SH  
 CC binding protein (SBP) receptor (I). The method involves providing an  
 CC assay for measuring binding to or uptake by (I) of SH bound to or  
 CC complexed with SBP, adding the compound to be tested to the assay, and  
 CC determining the amount of SH bound to or complexed with SBP which is  
 CC bound to or taken up by (I), comparing the amount determined with an  
 CC amount measured in the absence of the compound to be tested, where a  
 CC difference in the two amounts identifies a compound which alters the  
 CC binding or uptake of SH bound to or complexed with SBP. Also described is  
 CC a compound (II) which alters the uptake of SH into cells presenting (I)  
 CC as identified by (M1). (I) has cytostatic and contraceptive activities.  
 CC (M1) is useful for screening for a compound which alters uptake of SH  
 CC such as androgens, oestrogens, progesterone and corticoids, preferably

CC androgens and oestrogens (e.g. testosterone, dihydrotestosterone, and  
 CC oestradiol), into cells presenting (I). (II) is useful for treating a  
 CC clinical condition such as prostate or breast cancer in an individual.  
 CC (II) is also useful for preventing pregnancy in a male. (II) is useful  
 CC for preparing a pharmaceutical composition for treating a clinical  
 CC condition associated with a steroid hormone, e.g. prostate cancer or  
 CC breast cancer. The present sequence represents the specifically claimed  
 CC human low density lipoprotein receptor related protein 1B (LRP1B)  
 CC protein, which is given in the exemplification of the present invention  
 XX

SQ Sequence 4599 AA;

Query Match 55.3%; Score 280; DB 6; Length 4599;

Best Local Similarity 53.6%; Pred. No. 6.9e-19;

Matches 45; Conservative 12; Mismatches 27; Indels 0; Gaps 0;

QY 3 CSPKQACRQITICISKGWCDGRCPCDGSDEAPEICQSKAQRCPNEHNCIGTGLCV 62

Db 32 CDGFEFLCHDHVTCVQSQWLCDGDPDSDSDSLDTCPEEVEIKPLNHIACLGTKVCV 91

QY 63 PMSRLNGVQDCMDGSDGPHCRE 86

Db 92 HLSQLCNGVLDCPDGYDEGVHCQE 115

Search completed: September 17, 2004, 11:15:23

Job time : 124 secs

This Page Blank (uspto)

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2004, 11:13:19 ; Search time 32 Seconds  
(without alignments)  
138.745 Million cell updates/sec

Title: US-09-625-137-21

Perfect score: 506

Sequence: 1 KTCSPKQFACRDQITCSKG.....LCNGVQDCMGSDGPHCRE 86

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

1: /cgn2\_6/prodata/2/iaa/5A-COMB.pep.\*

2: /cgn2\_6/prodata/2/iaa/5B-COMB.pep.\*

3: /cgn2\_6/prodata/2/iaa/6A-COMB.pep.\*

4: /cgn2\_6/prodata/2/iaa/6B-COMB.pep.\*

5: /cgn2\_6/prodata/2/iaa/6CTUS-COMB.pep.\*

6: /cgn2\_6/prodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	506	100.0	4544	1	US-08-469-486-52
2	506	100.0	4544	2	US-08-469-658-52
3	202	39.9	726	6	5208144-37
4	191.5	37.8	159	6	5208144-35
5	187.5	37.1	2214	1	US-08-727-034-7
6	184.5	36.5	2213	1	US-08-727-034-3
7	183	36.2	4654	3	US-08-476-515A-84
8	183	36.2	4655	3	US-08-652-877-84
9	183	36.2	4655	3	US-08-652-877-88
10	183	36.2	4655	3	US-08-652-877-90
11	182	36.0	4655	3	US-08-652-877-86
12	180.5	35.7	846	1	US-08-149-103-3
13	180.5	35.7	846	1	US-08-149-103-4
14	180.5	35.7	846	1	US-08-451-883-3
15	180.5	35.7	846	1	US-08-451-883-4
16	180.5	35.7	873	1	US-08-393-734-2
17	180.5	35.7	873	1	US-08-894-489-2
18	176.5	34.9	884	6	5208144-8
19	176	34.8	356	1	US-08-228-162-2
20	176	34.8	860	1	US-08-092-817-4
21	176	34.8	860	4	US-08-485-128-4
22	176	34.8	860	4	US-08-804-778A-8
23	176	34.8	1074	3	US-08-470-058-2
24	176	34.8	1074	3	US-09-037-188-2
25	176	34.8	1074	3	US-09-285-310-2
26	176	34.8	1410	2	US-08-470-058-4
27	176	34.8	1410	3	US-09-037-188-4

28 176 34.8 1410 3 US-09-285-310-4 Sequence 4, Appli  
29 174.5 34.5 943 3 US-08-476-515A-12 Sequence 12, Appl  
30 174.5 34.5 944 3 US-08-652-877-12 Sequence 12, Appl  
31 171.5 33.9 1451 4 US-09-060-299-25 Sequence 25, Appl  
32 171.5 33.9 1451 4 US-09-402-923A-25 Sequence 25, Appl  
33 171.5 33.9 1584 4 US-09-060-299-39 Sequence 39, Appl  
34 171.5 33.9 1584 4 US-09-402-923A-39 Sequence 39, Appl  
35 171.5 33.9 1591 4 US-09-060-299-4 Sequence 4, Appl  
36 171.5 33.9 1591 4 US-09-060-299-43 Sequence 43, Appl  
37 171.5 33.9 1591 4 US-09-402-923A-4 Sequence 4, Appl  
38 171.5 33.9 1591 4 US-09-402-923A-43 Sequence 43, Appl  
39 171.5 33.9 1615 4 US-09-060-299-3 Sequence 3, Appl  
40 171.5 33.9 1615 4 US-09-402-923A-3 Sequence 3, Appl  
41 171.5 33.9 1639 4 US-09-060-299-29 Sequence 29, Appl  
42 171.5 33.9 1639 4 US-09-402-923A-29 Sequence 29, Appl  
43 170.5 33.7 1586 4 US-09-060-299-44 Sequence 44, Appl  
44 170.5 33.7 1586 4 US-09-402-923A-44 Sequence 44, Appl  
45 170.5 33.7 1614 4 US-09-060-299-42 Sequence 42, Appl

#### ALIGNMENTS

#### RESULT 1

US-08-469-486-52  
; Sequence 52, Application US/08469486  
; Patent No. 5739281  
; GENERAL INFORMATION:  
; APPLICANT: Thøgersen, Hans Christian  
; APPLICANT: Holtet, Thor Las  
; APPLICANT: Ezerodt, Michael  
; TITLE OF INVENTION: Improved method for the refolding of  
; TITLE OF INVENTION: proteins  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version  
; SOFTWARE: #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/469,486  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/192,060  
; FILING DATE: February 4, 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Paul T. Clark  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 06363/002001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617 542 5070  
; TELEFAX: 617 542 8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 52:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4544 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-469-486-52

Query Match 100.0%; Score 506; DB 1; Length 4544;  
Best Local Similarity 100.0%; Pred. No. 8.6e-45;

```
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCSPKQFACRDQITCISKWRCGDRDCPDGSDAEAPICPSKAQRCPQNEHNCIGTEL 60
DB 25 KTCSPKQFACRDQITCISKWRCGDRDCPDGSDAEAPICPSKAQRCPQNEHNCIGTEL 84
QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86
DB 85 CVPMSRLCNGVQDCMDGSDGPHCRE 110

RESULT 2
US-08-469-658-52
; Sequence 52, Application US/08469658
; Patent No. 5917018
; GENERAL INFORMATION:
; APPLICANT: Th egeresen, Hans Christian
; APPLICANT: Holtet, Thor Las
; APPLICANT: Etzerodt, Michael
; TITLE OF INVENTION: IMPROVED METHOD FOR THE REFOLDING OF
; PROTEINS
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,658
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/192,060
; FILING DATE: February 4, 1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul T. Clark
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06363/002002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 542 5070
; TELEFAX: 617 542 8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 454 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-469-658-52

Query Match 100.0%; Score 506; DB 2; Length 454;
Best Local Similarity 100.0%; Pred. No. 8.6e-45;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCSPKQFACRDQITCISKWRCGDRDCPDGSDAEAPICPSKAQRCPQNEHNCIGTEL 60
DB 25 KTCSPKQFACRDQITCISKWRCGDRDCPDGSDAEAPICPSKAQRCPQNEHNCIGTEL 84
QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86
DB 85 CVPMSRLCNGVQDCMDGSDGPHCRE 110

Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCSPKQFACRDQITCISKWRCGDRDCPDGSDAEAPICPSKAQRCPQNEHNCIGTEL 60
DB 25 KTCSPKQFACRDQITCISKWRCGDRDCPDGSDAEAPICPSKAQRCPQNEHNCIGTEL 84
QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86
DB 85 CVPMSRLCNGVQDCMDGSDGPHCRE 110
```

```
RESULT 3
5208144-37
; Patent No. 5208144
; APPLICANT: SMITH, JOHN A.;RAYCHOWDHURY, RAKTIMA;NILES,JOHN L.
; TITLE OF INVENTION: METHOD FOR DETECTION OF HUMAN DNA
; CONTAINING THE GENE ENCODING LOW DENSITY LIPOPROTEIN RECEPTOR
; NUMBER OF SEQUENCES: 42
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/396,697
; FILING DATE: 22-AUG-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 313,682
; FILING DATE: 22-FEB-1989
; APPLICATION NUMBER: 235,211
; FILING DATE: 23-AUG-1988
; SEQ ID NO:37:
; LENGTH: 726
5208144-37

Query Match 39.9%; Score 202; DB 6; Length 726;
Best Local Similarity 45.3%; Pred. No. 1.8e-13;
Matches 39; Conservative 11; Mismatches 32; Indels 4; Gaps 3;
QY 1 KTCSPKQFACRDQITCISKWRCGDRDCPDGSDAEAPICPSKAQRCPQNEHNCIGTEL 60
DB 397 KTCSPFTHFLC-DNGNCIYKAWICDGDNDRCRMSDE--KDCP-TQPFHCPSTQWQCPGYST 452
QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86
DB 453 CINLSALCDGVFCPCNGTDESPLCNQ 478

RESULT 4
5208144-35
; Patent No. 5208144
; APPLICANT: SMITH, JOHN A.;RAYCHOWDHURY, RAKTIMA;NILES,JOHN L.
; TITLE OF INVENTION: METHOD FOR DETECTION OF HUMAN DNA
; CONTAINING THE GENE ENCODING LOW DENSITY LIPOPROTEIN RECEPTOR
; NUMBER OF SEQUENCES: 42
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/396,697
; FILING DATE: 22-AUG-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 313,682
; FILING DATE: 22-FEB-1989
; APPLICATION NUMBER: 235,211
; FILING DATE: 23-AUG-1988
; SEQ ID NO:35:
; LENGTH: 159
5208144-35

Query Match 37.8%; Score 191.5; DB 6; Length 159;
Best Local Similarity 41.0%; Pred. No. 4.4e-13;
Matches 34; Conservative 10; Mismatches 36; Indels 3; Gaps 1;
QY 2 TCSPPKQFACRDQITCISKWRCGDRDCPDGSDAEAPICPSKAQRCPQNEHNCIGTEL 61
DB 2 TCPDPFTKQCTNTICVPRFLCDGNDGSDGSDENPIYC---ASHTCRSNEFQCLSFQRC 58
QY 62 VPMMSRLCNGVQDCMDGSDGPHCRE 84
DB 59 IPSYWFCDGEADCADGSDGSDGPHCRE 81

RESULT 5
US-08-727-034-7
; Sequence 7, Application US/08727034
; Patent No. 5655872
; GENERAL INFORMATION:
; APPLICANT: SAITO, YASHUSHI
; APPLICANT: IWASAKI, AKIO
; APPLICANT: ARAI, KOICHI
; APPLICANT: YAMAZAKI, HIROYUKI
```

; TITLE OF INVENTION: NOVEL LDL RECEPTOR ANALOG PROTEIN AND
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/727,034
; FILING DATE: 08-OCT-1996
; CLASSIFICATION: 536
; PRIORITY APPLICATION DATA:
; PRIOR APPLICATION NUMBER: JP 261440/1995
; FILING DATE: 09-OCT-1995
; APPLICATION NUMBER: JP 102451/1996
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 80-079-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2214 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-727-034-7
;
; Query Match 37.1%; Score 187.5; DB 1; Length 2214;
; Best Local Similarity 47.6%; Pred. No. 2.1e-11;
; Matches 40; Conservative 9; Mismatches 24; Indels 11; Gaps 6;
;
; QY 2 TCSPKQFACRDQITCTISKGRWCDGRCDEPDCGSDPAEIPCQS-KAQRCPNHNCLGTTEL 60
; Db 1198 TCEASNFQCRNG-HCIPQRWACDGTDCQDGSDEPVCCKKNGFRC-PN-----GT-- 1248
;
; QY 61 CVPMSRLCNGVQDCMGSDGSGPHC 84
; Db 1249 CIPSSKHCDGLRDCSDGSDSDE-QHC 1271
;
; RESULT 6
; US-08-727-034-3
; Sequence 3, Application US/08727034
; Patent No. 5665872
; GENERAL INFORMATION:
; APPLICANT: SAITO, YASHUSHI
; APPLICANT: IWASAKI, AKIO
; APPLICANT: ARAI, KOICHI
; APPLICANT: YAMAZAKI, HIROYUKI
; TITLE OF INVENTION: NOVEL LDL RECEPTOR ANALOG PROTEIN AND
; TITLE OF INVENTION: THE GENE CODING THEREFOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA

; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/727,034
; FILING DATE: 08-OCT-1996
; CLASSIFICATION: 536
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 261440/1995
; FILING DATE: 09-OCT-1995
; APPLICATION NUMBER: JP 102451/1996
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 80-079-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2213 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-727-034-3
;
; Query Match 36.5%; Score 184.5; DB 1; Length 2213;
; Best Local Similarity 46.4%; Pred. No. 4.4e-11;
; Matches 39; Conservative 9; Mismatches 25; Indels 11; Gaps 6;
;
; QY 2 TCSPKQFACRDQITCTISKGRWCDGRCDEPDCGSDPAEIPCQS-KAQRCPNHNCLGTTEL 60
; Db 1197 TCEASNFQCRNG-HCIPQRWACDGTDCQDGSDEPVCCKKNGFRC-PN-----GT-- 1247
;
; QY 61 CVPMSRLCNGVQDCMGSDGSGPHC 84
; Db 1248 CIPSTKHCDGLRDCSDGSDSDE-QHC 1270
;
; RESULT 7
; US-08-476-515A-84
; Sequence 84, Application US/08476515A
; Patent No. 6239270
; GENERAL INFORMATION:
; APPLICANT: Akerstrom, Goran
; APPLICANT: Tuhlin, Claes
; APPLICANT: Rask, Lars
; APPLICANT: Crumley, Gregg R.
; APPLICANT: Morse, Clarence C.
; APPLICANT: Murray, Edward M.
; APPLICANT: Hjaln, Goran
; TITLE OF INVENTION: Human Calcium Sensor Protein, Fragments
; TITLE OF INVENTION: Thersof and DNA Encoding Same
; NUMBER OF SEQUENCES: 84
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Martin Savitzky
; STREET: Rhone-Poulenc Rorer Inc., 500 Arcola Rd.;
; STREET: 3C43,
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426-0107
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Compaq PC
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Word 7.0 (Patentin)
; CURRENT APPLICATION DATA:

```

; APPLICATION NUMBER: US/08/476,515A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/344,836
; FILING DATE: 23-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,314
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Savitzky, Martin
; REGISTRATION NUMBER: 29,699
; REFERENCE/DOCKET NUMBER: A1355D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-454-3816
; TELEFAX: 610-454-3808
; INFORMATION FOR SEQ ID NO: 84:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4654 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-476-515A-84

Query Match 36.2%; Score 183; DB 3; Length 4654;
Best Local Similarity 42.9%; Pred. No. 1.4e-10;
Matches 36; Conservative 13; Mismatches 31; Indels 4; Gaps 3;

Qy 1 KTCSPKQFACDQITCISKGRCDGDCPDGSDPEAICPQSKAQCQPNHNCLGTGL 60
Db 1269 KTCPSYFHC-DNGNCIHRALWCLDRNDGDMDSDE--KDCP-TQPPRCPSWQWQCLGHNI 1324

Qy 61 CVPMSRLNGVQDCMDGSDGPHC 84
Db 1325 CVNLSVVCDGIFDCPNGTDESPLC 1348

RESULT 8
US-08-652-877-84
; Sequence 84, Application US/08652877
; Patent No. 6187548
; GENERAL INFORMATION:
; APPLICANT: Akerstrom, Goran
; APPLICANT: Juhlin, Claes
; APPLICANT: Rask, Lars
; APPLICANT: Crumley, Gregg R.
; APPLICANT: Morse, Clarence C.
; APPLICANT: Murray, Edward M.
; APPLICANT: Hjalms, Goran
; TITLE OF INVENTION: Human Calcium Sensor Protein, Fragments
; TITLE OF INVENTION: Thereof and DNA Encoding Same
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Rd., 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426-0107
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: System 7.5.1
; SOFTWARE: Word 6.0 (Patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,877
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/15203
; FILING DATE: 22-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/344,836
; FILING DATE: 23-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/15203

```

```

; FILING DATE: 22-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/344,836
; FILING DATE: 23-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,314
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Savitzky, Martin
; REGISTRATION NUMBER: 29,699
; REFERENCE/DOCKET NUMBER: A1355E-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-454-3816
; TELEFAX: 610-454-3808
; INFORMATION FOR SEQ ID NO: 84:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4655 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-652-877-84

Query Match 36.2%; Score 183; DB 3; Length 4655;
Best Local Similarity 42.9%; Pred. No. 1.4e-10;
Matches 36; Conservative 13; Mismatches 31; Indels 4; Gaps 3;

Qy 1 KTCSPKQFACDQITCISKGRCDGDCPDGSDPEAICPQSKAQCQPNHNCLGTGL 60
Db 1269 KTCPSYFHC-DNGNCIHRALWCLDRNDGDMDSDE--KDCP-TQPPRCPSWQWQCLGHNI 1324

Qy 61 CVPMSRLNGVQDCMDGSDGPHC 84
Db 1325 CVNLSVVCDGIFDCPNGTDESPLC 1348

RESULT 9
US-08-652-877-88
; Sequence 88, Application US/08652877
; Patent No. 6187548
; GENERAL INFORMATION:
; APPLICANT: Akerstrom, Goran
; APPLICANT: Juhlin, Claes
; APPLICANT: Rask, Lars
; APPLICANT: Crumley, Gregg R.
; APPLICANT: Morse, Clarence C.
; APPLICANT: Murray, Edward M.
; APPLICANT: Hjalms, Goran
; TITLE OF INVENTION: Human Calcium Sensor Protein, Fragments
; TITLE OF INVENTION: Thereof and DNA Encoding Same
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Rd., 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426-0107
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: System 7.5.1
; SOFTWARE: Word 6.0 (Patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,877
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/15203
; FILING DATE: 22-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/344,836
; FILING DATE: 23-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/15203

```

APPLICATION NUMBER: US 08/487,314  
FILING DATE: 07-JUNE-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Savitzky, Martin  
REGISTRATION NUMBER: 29,699  
REFERENCE/DOCKET NUMBER: A1355E-US  
TELEPHONE: 610-454-3816  
TELEFAX: 610-454-3808  
INFORMATION FOR SEQ ID NO: 88:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4655 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-652-877-88

Query Match 36.2%; Score 183; DB 3; Length 4655;  
Best Local Similarity 42.9%; Pred. No. 1.4e-10;  
Matches 36; Conservative 13; Mismatches 31; Indels 4; Gaps 3;

QY 1 KTCSPKQFACRDQITCISKGRCDGSDGDEAPETCPQSKAQCQPNHNCLGTGL 60  
DB 1269 KTCPSVFFHC-DNGNCIHRWLCDRDNDGDSDE--KDCP-TQPFRCPSWQWQCLGHNI 1324

QY 61 CVPMSRLCNGVQDCMDGSDGPHC 84  
DB 1325 CYNLSVVCDFGFCPNGTDESPLC 1348

RESULT 10  
US-08-652-877-90  
Sequence 90, Application US/08652877  
Patent No. 6187548  
GENERAL INFORMATION:  
APPLICANT: Akerstrom, Goran  
APPLICANT: Juhlin, Claes  
APPLICANT: Rask, Lars  
APPLICANT: Crumley, Gregg R.  
APPLICANT: Morse, Clarence C.  
APPLICANT: Murray, Edward M.  
APPLICANT: Hjaln, Goran  
TITLE OF INVENTION: Human Calcium Sensor Protein, Fragments  
TITLE OF INVENTION: Thereof and DNA Encoding Same  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcoia Rd., 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426-0107  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: System 7.5.1  
SOFTWARE: Word 6.0 (patentin)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/652,877  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/15203  
FILING DATE: 22-NOV-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/344,836  
FILING DATE: 23-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/487,314  
FILING DATE: 07-JUNE-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Savitzky, Martin  
REGISTRATION NUMBER: 29,699  
REFERENCE/DOCKET NUMBER: A1355E-US  
TELEPHONE: 610-454-3816  
TELEFAX: 610-454-3808  
INFORMATION FOR SEQ ID NO: 86:

REFERENCE/DOCKET NUMBER: A1355E-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-454-3816  
TELEFAX: 610-454-3808  
INFORMATION FOR SEQ ID NO: 90:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4655 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-652-877-90

Query Match 36.2%; Score 183; DB 3; Length 4655;  
Best Local Similarity 42.9%; Pred. No. 1.4e-10;  
Matches 36; Conservative 13; Mismatches 31; Indels 4; Gaps 3;

QY 1 KTCSPKQFACRDQITCISKGRCDGSDGDEAPETCPQSKAQCQPNHNCLGTGL 60  
DB 1269 KTCPSVFFHC-DNGNCIHRWLCDRDNDGDSDE--KDCP-TQPFRCPSWQWQCLGHNI 1324

QY 61 CVPMSRLCNGVQDCMDGSDGPHC 84  
DB 1325 CYNLSVVCDFGFCPNGTDESPLC 1348

RESULT 11  
US-08-652-877-86  
Sequence 86, Application US/08652877  
Patent No. 6187548  
GENERAL INFORMATION:  
APPLICANT: Akerstrom, Goran  
APPLICANT: Juhlin, Claes  
APPLICANT: Rask, Lars  
APPLICANT: Crumley, Gregg R.  
APPLICANT: Morse, Clarence C.  
APPLICANT: Murray, Edward M.  
APPLICANT: Hjaln, Goran  
TITLE OF INVENTION: Human Calcium Sensor Protein, Fragments  
TITLE OF INVENTION: Thereof and DNA Encoding Same  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcoia Rd., 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426-0107  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: System 7.5.1  
SOFTWARE: Word 6.0 (patentin)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/652,877  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/15203  
FILING DATE: 22-NOV-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/344,836  
FILING DATE: 23-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/487,314  
FILING DATE: 07-JUNE-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Savitzky, Martin  
REGISTRATION NUMBER: 29,699  
REFERENCE/DOCKET NUMBER: A1355E-US  
TELEPHONE: 610-454-3816  
TELEFAX: 610-454-3808  
INFORMATION FOR SEQ ID NO: 86:

```

;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4655 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-652-877-86

```

Query Match 36.0%; Score 182; DB 3; Length 4655;  
Best Local Similarity 44.3%; Pred. No. 1.8e-10;  
Matches 35; Conservative 14; Mismatches 24; Indels

QY	2	TCSPQAFACRQDITCISKHRCGDERDCDPSGDEAETCPSQAKAORCOPNEHNCJGTCLC	61
Db	1108	SCLDQTYTC-DNHQICISKHWWCDDNDGCGSDE--KNC--NSTETCQPSQFNC-PNHRC	1161
QY	62	VPMRLCNGVQCDCMGDSDE	80
Db	1162	IDL8FVCDGDKDCVDGSD	1180

RESULT 12  
US-8-149-103-3  
; Sequence 3, Application US/08149103  
; Patent No. 5750367  
; GENERAL INFORMATION:  
; APPLICANT: Lawrence C. B. Chan  
; TITLE OF INVENTION: HUMAN AND MOUSE VERY LOW  
; TITLE OF INVENTION: DENSITY LIPOPROTEIN RECEPTORS  
; TITLE OF INVENTION: AND METHODS FOR USE OF SUCH  
; TITLE OF INVENTION: RECEPTORS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LYON & LYON  
; STREET: 611 West Sixth Street  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90017

Query Match 35.7%; Score 180.5; DB 1; Length 846;  
Best Local Similarity 44.9%; Pred. NO. 4e-11;  
Matches 35; Conservative 11; Mismatches 25; Indels 7; Gaps 4;

QY 3 CSPKQFACRDQITCISKGWRCDGERDCPDGSDEAPEICPQSKAQRCPNEHNCGLTEL CV 62

[illegible]

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match 35.7%; Score 180.5; DB 1; Length 846;  
Best Local Similarity 44.9%; Pred. No. 4e-11;  
Matches 35: Conservative 11; Mismatches 25; Indels 7; Gaps 4;

Qy	3	CSPKQFACRQDITCISKWRCDCGERDCPDGSDDEAPEICQSKAQBCQPNHNCIGTELVCV	62
D6	212	CSTSEIQCGSG-ECIHKWRCDGDPCKDGSDEVN--CP---SRTCPRDGFECEDGS-CI	264

Qy 63 PMSRLCNGVQDCMDGSD 80  
 Db 265 HGSRCNGIRDCVDSDE 282

RESULT 14  
US-08-451-883-3  
; Sequence 3, Application US/08451883  
; Patent No. 5798209





Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2004, 11:15:30 ; Search time 48 Seconds  
(without alignments)  
575.364 Million cell updates/sec

Title: US-09-625-137-21  
Perfect score: 506  
Sequence: 1 KTCSPKQFACRDQITCISKG.....LNGVQDCMDGSDGPHCRE 86

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1342398 seqs, 32113274 residues  
Total number of hits satisfying chosen parameters: 1342398

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	506	100.0	86	12	US-09-750-972-21
2	506	100.0	2641	16	US-10-741-601-333
3	506	100.0	4183	12	US-10-087-192-672
4	506	100.0	4485	16	US-10-741-601-332
5	506	100.0	4544	16	US-10-464-368-68
6	506	100.0	4563	12	US-10-276-774-1723
7	486	96.0	3197	12	US-10-087-192-669
8	486	96.0	4545	9	US-09-873-403-2
9	486	96.0	4545	12	US-09-750-972-2
10	486	96.0	4545	16	US-10-464-368-67
11	486	96.0	4545	16	US-10-464-368-71
12	281	55.5	4636	9	US-09-835-996A-33
13	280	55.3	6339	9	US-09-835-996A-15
14	270	53.4	4599	16	US-10-464-368-69
15	270	53.4	4599	16	US-10-464-368-70

16	258	51.0	44	12	US-09-750-972-20	Sequence 20, Appl
17	253	50.0	43	12	US-09-750-972-22	Sequence 22, Appl
18	227	44.9	38	14	US-10-133-128-78	Sequence 78, Appl
19	227	44.9	38	14	US-10-289-660-78	Sequence 78, Appl
20	227	44.9	38	16	US-10-693-057-78	Sequence 79, Appl
21	224	44.3	37	14	US-10-133-128-79	Sequence 79, Appl
22	224	44.3	37	14	US-10-289-660-79	Sequence 79, Appl
23	224	44.3	37	16	US-10-693-057-79	Sequence 79, Appl
24	212	41.9	99	16	US-10-693-057-417	Sequence 417, Appl
25	211.5	41.8	81	12	US-09-750-972-52	Sequence 52, Appl
26	211.5	41.8	99	16	US-10-693-057-421	Sequence 421, Appl
27	211.5	41.8	123	12	US-09-750-972-50	Sequence 50, Appl
28	211.5	41.8	170	12	US-09-750-972-47	Sequence 47, Appl
29	211.5	41.8	209	12	US-09-750-972-44	Sequence 44, Appl
30	211.5	41.8	248	12	US-09-750-972-40	Sequence 40, Appl
31	211.5	41.8	289	12	US-09-750-972-35	Sequence 35, Appl
32	210	41.5	90	16	US-10-693-057-422	Sequence 422, Appl
33	204.5	40.4	4753	15	US-10-369-493-5119	Sequence 5119, Ap
34	202	39.9	4660	16	US-10-464-368-74	Sequence 74, Appl
35	201.5	39.8	90	16	US-10-693-057-425	Sequence 425, Appl
36	199	38.3	97	16	US-10-693-057-434	Sequence 434, Appl
37	194	38.3	819	15	US-10-094-749-1690	Sequence 1690, Ap
38	193	38.1	92	16	US-10-693-057-416	Sequence 416, Appl
39	193	38.1	963	16	US-10-464-368-86	Sequence 86, Appl
40	192.5	38.0	92	16	US-10-693-057-432	Sequence 432, Appl
41	191.5	37.8	2643	15	US-10-369-493-5010	Sequence 5010, Ap
42	190.5	37.6	862	14	US-10-281-478-3	Sequence 3, Appli
43	190.5	37.6	862	16	US-10-464-368-90	Sequence 90, Appl
44	190.5	37.6	862	16	US-10-464-368-91	Sequence 91, Appl
45	189	37.4	2180	15	US-10-369-493-5009	Sequence 5009, Ap

ALIGNMENTS

RESULT 1  
US-09-750-972-21  
; Sequence 21, Application US/09750972  
; Publiction No. US2004007293A1  
; GENERAL INFORMATION:  
; APPLICANT: Pramod K. Srivastava  
; TITLE OF INVENTION: ALPHA(2) MACROGLOBULIN RECEPTOR AS A HEAT SHOCK  
; FILE REFERENCE: 8449-134  
; CURRENT APPLICATION NUMBER: US/09/750,972  
; CURRENT FILING DATE: 2000-12-28  
; PRIOR APPLICATION NUMBER: 09/750,972  
; PRIOR FILING DATE: 2000-12-28  
; PRIOR APPLICATION NUMBER: 09/668,724  
; PRIOR FILING DATE: 2000-09-22  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 21  
; LENGTH: 86  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-750-972-21

Query Match	100.0%	Score 506,	DB 12;	Length 86;
Best Local Similarity	100.0%;	Pred. No. 2.1e-43;		
Matches	86;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	1	KTCSPKQFACRDQITCISKGWRCDGRCPCPDGSDGPHCRE	16	Sequence 60
Db	1	KTCSPKQFACRDQITCISKGWRCDGRCPCPDGSDGPHCRE	16	Sequence 60
QY	61	CVPMRLNGVQDCMDGSDGPHCRE	86	
Db	61	CVPMRLNGVQDCMDGSDGPHCRE	86	

RESULT 2  
US-10-741-601-333

```
; Sequence 333, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10741,601
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 333
; LENGTH: 2641
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-333

Query Match      100.0%; Score 506; DB 16; Length 2641;
Best Local Similarity 100.0%; Pred. No. 6.3e-42;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRDQITCISKWRCGDERDCPDGSDAPEICPQSKAQRCPNEHNCIGTEL 60
Db 25 KTCSPKQFACRDQITCISKWRCGDERDCPDGSDAPEICPQSKAQRCPNEHNCIGTEL 84

QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86
Db 85 CVPMSRLCNGVQDCMDGSDGPHCRE 110

RESULT 3
US-10-087-192-672
; Sequence 672, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 672
; LENGTH: 4183
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-087-192-672

Query Match      100.0%; Score 506; DB 12; Length 4183;
Best Local Similarity 100.0%; Pred. No. 1e-41;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRDQITCISKWRCGDERDCPDGSDAPEICPQSKAQRCPNEHNCIGTEL 60
Db 25 KTCSPKQFACRDQITCISKWRCGDERDCPDGSDAPEICPQSKAQRCPNEHNCIGTEL 84

QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86
Db 85 CVPMSRLCNGVQDCMDGSDGPHCRE 110

RESULT 4
US-10-741-601-332
; Sequence 332, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
```

```
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 332
; LENGTH: 4485
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-332

Query Match      100.0%; Score 506; DB 16; Length 4485;
Best Local Similarity 100.0%; Pred. No. 1.1e-41;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRDQITCISKWRCGDERDCPDGSDAPEICPQSKAQRCPNEHNCIGTEL 60
Db 25 KTCSPKQFACRDQITCISKWRCGDERDCPDGSDAPEICPQSKAQRCPNEHNCIGTEL 84

QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86
Db 85 CVPMSRLCNGVQDCMDGSDGPHCRE 110

RESULT 5
US-10-464-368-68
; Sequence 68, Application US/10464368
; Publication No. US20040023356A1
; GENERAL INFORMATION:
; APPLICANT: Krumlauf, Robb
; APPLICANT: Ellies, Debra
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATING BONE DEPOSITION
; FILE REFERENCE: 40716-IP-017
; CURRENT APPLICATION NUMBER: US/10/464,368
; CURRENT FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/388,970
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 68
; LENGTH: 4544
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-10-464-368-68

Query Match      100.0%; Score 506; DB 16; Length 4544;
Best Local Similarity 100.0%; Pred. No. 1.1e-41;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRDQITCISKWRCGDERDCPDGSDAPEICPQSKAQRCPNEHNCIGTEL 60
Db 25 KTCSPKQFACRDQITCISKWRCGDERDCPDGSDAPEICPQSKAQRCPNEHNCIGTEL 84

QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86
Db 85 CVPMSRLCNGVQDCMDGSDGPHCRE 110

RESULT 6
US-10-276-774-1723
; Sequence 1723, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245a1e1 Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
```

; PRIOR APPLICATION NUMBER: 09/496,914  
; PRIOR FILING DATE: 2000-02-03  
; NUMBER OF SEQ ID NOS: 2700  
; SOFTWARE: Custom  
; SEQ ID NO 1723  
; LENGTH: 4563  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-276-774-1723

Query Match 100.0%; Score 506; DB 12; Length 4563;  
Best Local Similarity 100.0%; Pred. No. 1.1e-41;  
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCSPKQFACRDQITCISKWRCDCGDCPDGSDAEAPICPQSKAQRCPNHNCLGTEL 60  
DB 44 KTCSPKQFACRDQITCISKWRCDCGDCPDGSDAEAPICPQSKAQRCPNHNCLGTEL 103  
QY 61 CVPMSRLCNGVQDCMDGSDGEGPHCRE 86  
DB 104 CVPMSRLCNGVQDCMDGSDGEGPHCRE 129

RESULT 7  
US-10-087-192-669  
; Sequence 669, Application US/10087192  
; Publication No. US20020182586A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David W.  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
; FILE REFERENCE: 52945200122  
; CURRENT APPLICATION NUMBER: US/10/087,192  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: US 09/747,377  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/798,586  
; PRIOR FILING DATE: 2001-03-02  
; NUMBER OF SEQ ID NOS: 2059  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 669  
; LENGTH: 3197  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-087-192-669

Query Match 96.0%; Score 486; DB 12; Length 3197;  
Best Local Similarity 95.3%; Pred. No. 7.8e-40;  
Matches 82; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 1 KTCSPKQFACRDQITCISKWRCDCGDCPDGSDAEAPICPQSKAQRCPNHNCLGTEL 60  
DB 26 KTCSPKQFACRDQITCISKWRCDCGDCPDGSDAEAPICPQSKAQRCPNHNCLGTEL 85  
QY 61 CVPMSRLCNGVQDCMDGSDGEGPHCRE 86  
DB 86 CVPMSRLCNGVQDCMDGSDGEGPHCRE 111

RESULT 8  
US-09-873-403-2  
; Sequence 2, Application US/09873403  
; Patent No. US2002028207A1  
; GENERAL INFORMATION:  
; APPLICANT: Srivastava, Pramod K  
; TITLE OF INVENTION: COMPLEXES OF ALPHA (2) MACROGLOBULIN AND ANTIGENIC  
; FILE REFERENCE: 8449-178  
; CURRENT APPLICATION NUMBER: US/09/873,403  
; PRIOR FILING DATE: 2001-06-04  
; PRIOR APPLICATION NUMBER: 09/625,139  
; PRIOR FILING DATE: 2000-07-25

; PRIOR APPLICATION NUMBER: 60/209,266  
; PRIOR FILING DATE: 2000-06-02  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 4545  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-873-403-2

Query Match 96.0%; Score 486; DB 9; Length 4545;  
Best Local Similarity 95.3%; Pred. No. 1.1e-39;  
Matches 82; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 1 KTCSPKQFACRDQITCISKWRCDCGDCPDGSDAEAPICPQSKAQRCPNHNCLGTEL 60  
DB 26 KTCSPKQFACRDQITCISKWRCDCGDCPDGSDAEAPICPQSKAQRCPNHNCLGTEL 85  
QY 61 CVPMSRLCNGVQDCMDGSDGEGPHCRE 86  
DB 86 CVPMSRLCNGVQDCMDGSDGEGPHCRE 111

RESULT 9  
US-09-750-972-2  
; Sequence 2, Application US/09750972  
; Publication No. US20040072993A1  
; GENERAL INFORMATION:  
; APPLICANT: Pramod K. Srivastava  
; TITLE OF INVENTION: ALPHA(2) MACROGLOBULIN RECEPTOR AS A HEAT SHOCK  
; FILE REFERENCE: 8449-134  
; CURRENT APPLICATION NUMBER: US/09/750,972  
; PRIOR FILING DATE: 2000-12-28  
; PRIOR APPLICATION NUMBER: 09/750,972  
; PRIOR FILING DATE: 2000-12-28  
; PRIOR APPLICATION NUMBER: 09/668,724  
; PRIOR FILING DATE: 2000-09-22  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 4545  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-750-972-2

Query Match 96.0%; Score 486; DB 12; Length 4545;  
Best Local Similarity 95.3%; Pred. No. 1.1e-39;  
Matches 82; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 1 KTCSPKQFACRDQITCISKWRCDCGDCPDGSDAEAPICPQSKAQRCPNHNCLGTEL 60  
DB 26 KTCSPKQFACRDQITCISKWRCDCGDCPDGSDAEAPICPQSKAQRCPNHNCLGTEL 85  
QY 61 CVPMSRLCNGVQDCMDGSDGEGPHCRE 86  
DB 86 CVPMSRLCNGVQDCMDGSDGEGPHCRE 111

RESULT 10  
US-10-464-368-67  
; Sequence 67, Application US/10464368  
; Publication No. US20040023356A1  
; GENERAL INFORMATION:  
; APPLICANT: Krumlauf, Robb  
; APPLICANT: Ellies, Debra  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATING BONE DEPOSITION  
; FILE REFERENCE: 40716-1P-017  
; CURRENT APPLICATION NUMBER: US/10/464,368  
; CURRENT FILING DATE: 2003-06-16  
; PRIOR APPLICATION NUMBER: 60/388,970  
; PRIOR FILING DATE: 2002-06-14  
; NUMBER OF SEQ ID NOS: 140

```
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 67
; LENGTH: 4545
; TYPE: PRT
; ORGANISM: MOUSE
US-10-464-368-67

Query Match          96.0%; Score 486; DB 16; Length 4545;
Best Local Similarity 95.3%; Pred. No. 1.1e-39;
Matches 82; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRDQITCISKWRCDCGERDCPDGSDPEAPECQSKAQRCPQNEHNCIGTGL 60
DB 26 KTCSPKQFACRDQITCISKWRCDCGERDCPDGSDPEAPECQSKAQRCPQNEHNCIGTGL 85

QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86
DB 86 CVPMSRLCNGIQDCMDGSDGAHCRE 111

RESULT 11
US-10-464-368-71
; Sequence 71, Application US/10464368
; Publication No. US20040023356A1
; GENERAL INFORMATION:
; APPLICANT: Krumlauf, Robb
; APPLICANT: Ellies, Debra
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATING BONE DEPOSITION
; FILE REFERENCE: 40718-IP-017
; CURRENT APPLICATION NUMBER: US/10/464,368
; CURRENT FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/388,970
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 71
; LENGTH: 4545
; TYPE: PRT
; ORGANISM: MOUSE
US-10-464-368-71

Query Match          96.0%; Score 486; DB 16; Length 4545;
Best Local Similarity 95.3%; Pred. No. 1.1e-39;
Matches 82; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRDQITCISKWRCDCGERDCPDGSDPEAPECQSKAQRCPQNEHNCIGTGL 60
DB 26 KTCSPKQFACRDQITCISKWRCDCGERDCPDGSDPEAPECQSKAQRCPQNEHNCIGTGL 85

QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86
DB 86 CVPMSRLCNGIQDCMDGSDGAHCRE 111

RESULT 12
US-09-835-996A-33
; Sequence 33, Application US/09835996A
; Patent No. US20020142953A1
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis
; APPLICANT: Loeb, Debra
; APPLICANT: Montgomery, Julie
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhao, Qing
; APPLICANT: Wehrman, Tom
; APPLICANT: Drmanac, Radoje
; APPLICANT: Ren, Feiyan
; APPLICANT: Qian, Xiaohong
; APPLICANT: Wang, Dunrui
```

```
; TITLE OF INVENTION: MATERIALS AND METHODS RELATING TO LIPID METABOLISM
; FILE REFERENCE: 28110/35915A
; CURRENT APPLICATION NUMBER: US/09/835,996A
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: US 60/197,137
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: US 09/714,936
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: US 09/667,298
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: US 09/667,298
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 09/598,042
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 33
; LENGTH: 4636
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-835-996A-33

Query Match          55.5%; Score 281; DB 9; Length 4636;
Best Local Similarity 53.6%; Pred. No. 4.4e-19;
Matches 45; Conservative 12; Mismatches 27; Indels 0; Gaps 0;

QY 3 CSPKQFACRDQITCISKWRCDCGERDCPDGSDPEAPECQSKAQRCPQNEHNCIGTGLCV 62
DB 69 CDPGEFLCHDHTVCVSRSLWLCDDGPDGSDPEAPECQSKAQRCPQNEHNCIGTGLCV 128

QY 63 PMSRLCNGVQDCMDGSDGPHCRE 86
DB 129 HLSQLCNGVLCDDGPDGSDGPHCRE 152

RESULT 13
US-09-835-996A-15
; Sequence 15, Application US/09835996A
; Patent No. US20020142953A1
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis
; APPLICANT: Loeb, Debra
; APPLICANT: Montgomery, Julie
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhao, Qing
; APPLICANT: Wehrman, Tom
; APPLICANT: Drmanac, Radoje
; APPLICANT: Ren, Feiyan
; APPLICANT: Qian, Xiaohong
; APPLICANT: Wang, Dunrui
; TITLE OF INVENTION: MATERIALS AND METHODS RELATING TO LIPID METABOLISM
; FILE REFERENCE: 28110/35915A
; CURRENT APPLICATION NUMBER: US/09/835,996A
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: US 60/197,137
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: US 09/714,936
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: US 09/667,298
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 09/598,042
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 639
; TYPE: PRT
```

Search completed: September 17, 2004, 11:20:08  
Job time : 49 secs

```

US-10-464-368-69
; Sequence 69, Application US/10464368
; Publication No. US20040023356A1
; GENERAL INFORMATION:
; APPLICANT: Krumlauf, Debra
; APPLICANT: Elles, Debra
; TITLE OF INVENTION, COMPOSITIONS AND METHODS FOR REGULATING BONE DEPOSITION
; FILE REFERENCE: 40716-IP-017
; CURRENT APPLICATION NUMBER: US/10/464,368
; CURRENT FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/388,970
; PRIOR FILING DATE: 2003-06-14
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 69
; LENGTH: 4599
; TYPE: PRT
; ORGANISM: MOUSE
US-10-464-368-69

```

```

RESULT 15
US-10-464-368-70
; Sequence 70, Application US/10464368
; Publication No. US2004002356A1
; GENERAL INFORMATION:
; APPLICANT: Krumlauf, Robb
; APPLICANT: Ellies, Debra
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATING BONE DEPOSITION
; FILE REFERENCE: 40716-1P-017
; CURRENT APPLICATION NUMBER: US/10/464,368
; CURRENT FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/388,970
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 70
; LENGTH: 4599
; TYPE: PRT
; ORGANISM: MOUSE
; US-10-464-368-70

```

Query Match	53.4%;	Score 270;	DB 16;	Length 4599;
Best Local Similarity	51.2%;	Pred. No. 5.6e-18;		

This Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2004, 11:11:04 ; Search time 39 Seconds  
(without alignments)  
212.115 Million cell updates/sec

Title: US-09-625-137-21  
Perfect score: 506  
Sequence: 1 KTCSPKQFACRDQITCISKG.....LCNGVQDCMGSDGPHCKRE 86

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78:\*

1: Pirl:\*

2: Pirl2:\*

3: Pirl3:\*

4: Pirl4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	506	100.0	4544	1 S02392	alpha-2-macroglobu
2	486	96.0	4545	1 S25111	alpha-2-macroglobu
3	459	90.7	4543	1 A53102	alpha-2-macroglobu
4	204.5	40.4	4753	1 A47437	LDL-receptor-relat
5	202	39.9	4660	2 T42737	gp330 protein prec
6	200.5	39.6	909	1 QRXL22	LDL receptor 2 pre
7	197.5	39.0	990	2 H88733	protein F32E10.3 l
8	197.5	39.0	1984	2 T13171	probable vitelloge
9	196	38.7	972	2 A30363	glycoprotein Gp330
10	192.5	38.0	909	1 QRXL21	LDL receptor 1 pre
11	191.5	37.8	879	1 QRRTLD	LDL receptor precu
12	191.5	37.8	2643	2 T29149	hypothetical prote
13	190.5	37.6	862	1 QRMSLD	LDL receptor precu
14	189	37.4	854	1 QRHYLD	LDL receptor precu
15	189	37.4	2180	2 T29764	hypothetical prote
16	187.5	37.1	1650	2 S53457	dominant autoantig
17	187	37.0	996	2 JE0237	apolipoprotein E r
18	185.5	36.7	869	1 JC4858	VLDL receptor prec
19	185	36.6	1160	2 F88369	protein unc-52 (im
20	185	36.6	2295	2 C88369	protein unc-52 (im
21	185	36.6	3375	2 T19821	hypothetical prote
22	184.5	36.5	873	2 T14852	VLDL receptor prec
23	182.5	36.1	837	1 A29512	LDL receptor precu
24	182.5	36.1	863	1 S51789	VLDL receptor prec
25	182.5	36.1	873	1 QRRTVD	VLDL receptor prec
26	180.5	35.7	873	1 A49729	LDL receptor prec
27	177.5	35.1	2215	2 T00348	LR11 protein - mou
28	176	34.8	860	1 QRHULD	LDL receptor precu
29	176	34.8	1357	2 T16860	hypothetical prote

RESULT 1

S02392

alpha-2-macroglobulin receptor precursor - human

N:Alternate names: CD91; LDL receptor-related protein 1; low density lipoprotein recept

C:Species: Homo sapiens (man)

C>Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 22-Jun-1999

C:Accession: S02392; S30027; I37998; A39210; S12538

R:Herz, J.; Hamann, U.; Rogne, S.; Myklebost, O.; Gausepohl, H.; Stanley, K.K.

EVBO J. 7, 4119-4127, 1988

A>Title: Surface location and high affinity for calcium of a 500-kd liver membrane prot

A:Reference number: S02392; MUID:89210795; PMID:3266596

A:Accession: S02392

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-4544 <HER>

A:Cross-references: EMBL:X13916; NID:G34338; PIDN:CAA32112.1; PID:G34339

R:Kristensen, T.

submitted to the EMBL Data Library, October 1990

A:Reference number: S30027

A:Accession: S30027

A:Molecule type: mRNA

A:Residues: 3275-3864 <KRI>

A:Cross-references: EMBL:X55077

R:Herz, J.; Kowal, R.C.; Goldstein, J.L.; Brown, M.S.

EMBO J. 9, 1769-1776, 1990

A>Title: Proteolytic processing of the 600 kd low density lipoprotein receptor-related

A:Reference number: S12538; MUID:90269210; PMID:2112085

A:Contents: annotation; site of proteolytic cleavage

R:Kurt, H.; Herz, J.; Stanley, K.K.

Biochim. Biophys. Acta 1009, 229-236, 1989

A>Title: Structure of the low-density lipoprotein receptor-related protein (LRP) promot

A:Reference number: I37998; MUID:90089395; PMID:2597675

A:Accession: I37998

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-11 <RES>

A:Cross-references: EMBL:X15424; NID:G34408; PIDN:CAA33464.1; PID:G34409

R:Strickland, D.K.; Ashcom, J.D.; Williams, S.; Burgess, W.H.; Migliorini, M.; Argraves

J. Biol. Chem. 265, 17401-17404, 1990

A>Title: Sequence identity between the alpha2-macroglobulin receptor and low density li

A:Reference number: A39210; MUID:91009181; PMID:1698775

A:Accession: A39210

A>Status: preliminary

A:Molecule type: protein

A:Residues: 150-166;234-238, 'X', 240-245, 'X', 247-252, 'G', 686-695;902-916;1096-1109, 'S', 'I'

C:Genetics:

A:Gene: GDB:LRP1; APR; LRP; A2MR

A:Cross-references: GDB:I19694; OMIM:107770

A:Map position: 12q13.1-12q13.3

C:Complex: The alpha-2-macroglobulin receptor complex consists of noncovalently-associat

ciated protein (see PIR:A39875).

C:Superfamily: alpha-2-macroglobulin receptor; EGF homology; LDL receptor ligand-binding;

C;Keywords: beta-hydroxyasparagine; beta-hydroxyaspartic acid; calcium binding; glycoprotein; 19/Domain: signal sequence #status predicted <SIG>  
 F;1-19/Domain: signal sequence #status predicted <SIG>  
 F;20-3943/Product: alpha-2-macroglobulin receptor 85K chain #status predicted <S15K>  
 F;27-64/Domain: LDL receptor ligand-binding repeat homology <LDL1>  
 F;72-108/Domain: LDL receptor ligand-binding repeat homology <LDL2>  
 F;115-148/Domain: EGF homology <EG1>  
 F;154-188/Domain: EGF homology <EG2>  
 F;198-239/Domain: LDL receptor WYTD-containing repeat homology <YW01>  
 F;240-381/Domain: LDL receptor WYTD-containing repeat homology <YW02>  
 F;292-334/Domain: LDL receptor WYTD-containing repeat homology <YW03>  
 F;335-378/Domain: LDL receptor WYTD-containing repeat homology <YW04>  
 F;379-420/Domain: LDL receptor WYTD-containing repeat homology <YW05>  
 F;421-468/Domain: LDL receptor WYTD-containing repeat homology <YW06>  
 F;478-519/Domain: EGF homology <EG3>  
 F;571-613/Domain: LDL receptor WYTD-containing repeat homology <YW07>  
 F;614-659/Domain: LDL receptor WYTD-containing repeat homology <YW08>  
 F;660-710/Domain: LDL receptor WYTD-containing repeat homology <YW09>  
 F;711-752/Domain: LDL receptor WYTD-containing repeat homology <YW10>  
 F;753-799/Domain: LDL receptor WYTD-containing repeat homology <YW11>  
 F;807-842/Domain: EGF homology <EG4>  
 F;854-890/Domain: LDL receptor ligand-binding repeat homology <LDL3>  
 F;895-931/Domain: LDL receptor ligand-binding repeat homology <LDL4>  
 F;936-971/Domain: LDL receptor ligand-binding repeat homology <LDL5>  
 F;976-1011/Domain: LDL receptor ligand-binding repeat homology <LDL6>  
 F;1015-1051/Domain: LDL receptor ligand-binding repeat homology <LDL7>  
 F;1062-1097/Domain: LDL receptor ligand-binding repeat homology <LDL8>  
 F;1104-1140/Domain: LDL receptor ligand-binding repeat homology <LDL9>  
 F;1145-1182/Domain: LDL receptor ligand-binding repeat homology <LDL10>  
 F;1185-1221/Domain: EGF homology <EG5>  
 F;1227-1261/Domain: EGF homology <EG6>  
 F;1269-1308/Domain: LDL receptor WYTD-containing repeat homology <YW12>  
 F;1309-1355/Domain: LDL receptor WYTD-containing repeat homology <YW13>  
 F;1356-1398/Domain: LDL receptor WYTD-containing repeat homology <YW14>  
 F;1399-1445/Domain: LDL receptor WYTD-containing repeat homology <YW15>  
 F;1446-1488/Domain: LDL receptor WYTD-containing repeat homology <YW16>  
 F;1489-1531/Domain: LDL receptor WYTD-containing repeat homology <YW17>  
 F;1540-1578/Domain: EGF homology <EG7>  
 F;1583-1626/Domain: LDL receptor WYTD-containing repeat homology <YW18>  
 F;1627-1669/Domain: LDL receptor WYTD-containing repeat homology <YW19>  
 F;1670-1713/Domain: LDL receptor WYTD-containing repeat homology <YW20>  
 F;1714-1753/Domain: LDL receptor WYTD-containing repeat homology <YW21>  
 F;1754-1796/Domain: LDL receptor WYTD-containing repeat homology <YW22>  
 F;1797-1846/Domain: LDL receptor WYTD-containing repeat homology <YW23>  
 F;1850-1886/Domain: EGF homology <EG8>  
 F;1934-1976/Domain: LDL receptor WYTD-containing repeat homology <YW24>  
 F;1977-2019/Domain: LDL receptor WYTD-containing repeat homology <YW25>  
 F;2020-2063/Domain: LDL receptor WYTD-containing repeat homology <YW26>  
 F;2064-2105/Domain: LDL receptor WYTD-containing repeat homology <YW27>  
 F;2106-2151/Domain: LDL receptor WYTD-containing repeat homology <YW28>  
 F;2159-2194/Domain: EGF homology <EG9>  
 F;2199-2241/Domain: LDL receptor WYTD-containing repeat homology <YW29>  
 F;2253-2294/Domain: LDL receptor WYTD-containing repeat homology <YW30>  
 F;2344-2388/Domain: LDL receptor WYTD-containing repeat homology <YW31>  
 F;2389-2429/Domain: LDL receptor WYTD-containing repeat homology <YW32>  
 F;2430-2473/Domain: LDL receptor WYTD-containing repeat homology <YW33>  
 F;2482-2517/Domain: EGF homology <EG10>  
 F;2524-2561/Domain: LDL receptor ligand-binding repeat homology <LDL8>  
 F;2566-2600/Domain: LDL receptor ligand-binding repeat homology <LDL9>  
 F;2605-2639/Domain: LDL receptor ligand-binding repeat homology <LDL10>  
 F;2652-2688/Domain: LDL receptor ligand-binding repeat homology <LDL11>  
 F;2696-2730/Domain: LDL receptor ligand-binding repeat homology <LDL12>  
 F;2734-2769/Domain: LDL receptor ligand-binding repeat homology <LDL13>  
 F;2774-2812/Domain: LDL receptor ligand-binding repeat homology <LDL14>  
 F;2818-2853/Domain: LDL receptor ligand-binding repeat homology <LDL15>  
 F;2858-2897/Domain: LDL receptor ligand-binding repeat homology <LDL16>  
 F;2904-2939/Domain: LDL receptor ligand-binding repeat homology <LDL17>  
 F;2944-2980/Domain: EGF homology <EG11>  
 F;2986-3021/Domain: EGF homology <EG12>  
 F;3029-3068/Domain: LDL receptor WYTD-containing repeat homology <YW34>  
 F;3069-3113/Domain: LDL receptor WYTD-containing repeat homology <YW35>  
 F;3114-3156/Domain: LDL receptor WYTD-containing repeat homology <YW36>  
 F;3157-3200/Domain: LDL receptor WYTD-containing repeat homology <YW37>  
 F;3201-3241/Domain: LDL receptor WYTD-containing repeat homology <YW38>

F;3242-3284/Domain: LDL receptor WYTD-containing repeat homology <YW39>  
 F;3294-3330/Domain: EGF homology <EG13>  
 F;3334-3369/Domain: LDL receptor ligand-binding repeat homology <LDL1>  
 F;3374-3408/Domain: LDL receptor ligand-binding repeat homology <LDL2>  
 F;3413-3448/Domain: LDL receptor ligand-binding repeat homology <LDL3>  
 F;3453-3489/Domain: LDL receptor ligand-binding repeat homology <LDL4>  
 F;3494-3531/Domain: LDL receptor ligand-binding repeat homology <LDL5>  
 F;3536-3570/Domain: LDL receptor ligand-binding repeat homology <LDL6>  
 F;3573-3609/Domain: LDL receptor ligand-binding repeat homology <LDL7>  
 F;3613-3647/Domain: LDL receptor ligand-binding repeat homology <LDL8>  
 F;3654-3690/Domain: LDL receptor ligand-binding repeat homology <LDL9>  
 F;3695-3731/Domain: LDL receptor ligand-binding repeat homology <LDL10>  
 F;3741-3776/Domain: LDL receptor ligand-binding repeat homology <LDL11>  
 F;3785-3822/Domain: EGF homology <EG14>  
 F;3828-3860/Domain: EGF homology <EG15>  
 F;3868-3911/Domain: LDL receptor WYTD-containing repeat homology <YW40>  
 F;3912-3969/Domain: LDL receptor WYTD-containing repeat homology <YW41>  
 F;3944-4544/Product: alpha-2-macroglobulin receptor 85K chain #status predicted <85K>  
 F;3944-4420/Domain: 85K chain extracellular #status predicted <EXT>  
 F;3970-4012/Domain: LDL receptor WYTD-containing repeat homology <YW42>  
 F;4013-4056/Domain: LDL receptor WYTD-containing repeat homology <YW43>  
 F;4057-4099/Domain: LDL receptor WYTD-containing repeat homology <YW44>  
 F;4100-4142/Domain: LDL receptor WYTD-containing repeat homology <YW45>  
 F;4151-4182/Domain: EGF homology <EG16>  
 F;4200-4231/Domain: EGF homology <EG17>  
 F;4236-4267/Domain: EGF homology <EG18>  
 F;4272-4303/Domain: EGF homology <EG19>  
 F;4308-4339/Domain: EGF homology <EG20>  
 F;4344-4374/Domain: EGF homology <EG21>  
 F;4377-4408/Domain: EGF homology <EG22>  
 F;4421-4444/Domain: transmembrane #status predicted <TMM>  
 F;4445-4544/Domain: intracellular #status predicted <INT>  
 F;166-2598/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted  
 F;2598/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted  
 F;4075,4125,4278/Binding site: carbonylate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 506; DB 1; Length 4544;

Best Local Similarity 100.0%; Pred. No. 4.6e-38;  
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTSCKPQFACEDQITCTISKWRCDEGDCPDGSDPEICQSKACRCOPNEHNCIGTEL 60

Db 25 KTSCKPQFACRQITCTISKWRCDEGDCPDGSDPEICQSKACRCOPNEHNCIGTEL 84

QY 61 CVPMSRLCNGVQDCMDGSDGPHCRE 86

Db 85 CVPMSRLCNGVQDCMDGSDGPHCRE 110

# RESULT 2

S25111

alpha-2-macroglobulin receptor precursor - mouse

N:Alternate names: CD91; LDL receptor-related protein 1; low density lipoprotein receptor  
 C:Species: Mus musculus (house mouse)

C:Date: 04-Sep-1998 #sequence\_revision 04-Sep-1998 #text\_change 22-Jun-1999

C:Accession: S25111; S32554

R:van Leuven, F.

submitted to the EMBL Data Library, July 1992

A:Reference number: S25111

A:Accession: S25111

A:Molecule type: mRNA

A:Residues: 1-4545 <VANI>

A:Cross-references: EMBL:X67469; MID:g49941; PID:CAA47817.1; PID:g49942

R:van Leuven, F.; Stas, L.; Raymakers, L.; Overbergh, L.; de Strooper, B.; Halliker, C.;

Biochim. Biophys. Acta 1173, 71-74, 1993

A:Title: Molecular cloning and sequencing of the murine alpha-2-macroglobulin receptor c

A:Reference number: S32554; MUID:93250049; PMID:8485155

A:Accession: S32554

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-28;4416-4453 <VANI>

A:Cross-references: EMBL:X67469

C:Complex: The alpha-2-macroglobulin receptor complex consists of noncovalently-associated

ciated protein (see PIR:JX0281).

C/Superfamily: alpha-2-macroglobulin receptor; EGF homology; LDL receptor ligand-binding  
C/Keywords: beta-hydroxyasparagine; beta-hydroxyaspartic acid; calcium binding; glycopro  
F/1-19/Domain: signal sequence #status predicted <SIG>  
F/20-3944,3945-4545/Product: alpha-2-macroglobulin receptor #status predicted <MAT>  
F/20-3944/Domain: alpha-2-macroglobulin receptor 515K chain #status predicted <515K>  
F/28-65/Domain: LDL receptor ligand-binding repeat homology <LDL1>  
F/73-109/Domain: LDL receptor ligand-binding repeat homology <LDL2>  
F/116-149/Domain: EGF homology <EG1>  
F/155-189/Domain: EGF homology <EG2>  
F/199-240/Domain: LDL receptor WYTD-containing repeat homology <YW01>  
F/241-282/Domain: LDL receptor WYTD-containing repeat homology <YW02>  
F/293-335/Domain: LDL receptor WYTD-containing repeat homology <YW03>  
F/336-379/Domain: LDL receptor WYTD-containing repeat homology <YW04>  
F/380-421/Domain: LDL receptor WYTD-containing repeat homology <YW05>  
F/422-469/Domain: LDL receptor WYTD-containing repeat homology <YW06>  
F/479-520/Domain: EGF homology <EG3>  
F/572-614/Domain: LDL receptor WYTD-containing repeat homology <YW07>  
F/615-660/Domain: LDL receptor WYTD-containing repeat homology <YW08>  
F/661-711/Domain: LDL receptor WYTD-containing repeat homology <YW09>  
F/712-753/Domain: LDL receptor WYTD-containing repeat homology <YW10>  
F/754-800/Domain: LDL receptor WYTD-containing repeat homology <YW11>  
F/808-843/Domain: EGF homology <EG4>  
F/855-891/Domain: LDL receptor ligand-binding repeat homology <LDL3>  
F/896-932/Domain: LDL receptor ligand-binding repeat homology <LDL4>  
F/937-972/Domain: LDL receptor ligand-binding repeat homology <LDL5>  
F/977-1012/Domain: LDL receptor ligand-binding repeat homology <LDL6>  
F/1016-1052/Domain: LDL receptor ligand-binding repeat homology <LDL7>  
F/1063-1098/Domain: LDL receptor ligand-binding repeat homology <LDL8>  
F/1105-1141/Domain: LDL receptor ligand-binding repeat homology <LDL9>  
F/1146-1183/Domain: LDL receptor ligand-binding repeat homology <LDLA>  
F/1186-1222/Domain: EGF homology <EG5>  
F/1228-1262/Domain: EGF homology <EG6>  
F/1270-1309/Domain: LDL receptor WYTD-containing repeat homology <YW12>  
F/1310-1356/Domain: LDL receptor WYTD-containing repeat homology <YW13>  
F/1357-1399/Domain: LDL receptor WYTD-containing repeat homology <YW14>  
F/1400-1446/Domain: LDL receptor WYTD-containing repeat homology <YW15>  
F/1447-1489/Domain: LDL receptor WYTD-containing repeat homology <YW16>  
F/1490-1532/Domain: LDL receptor WYTD-containing repeat homology <YW17>  
F/1541-1579/Domain: EGF homology <EG7>  
F/1584-1627/Domain: LDL receptor WYTD-containing repeat homology <YW18>  
F/1628-1670/Domain: LDL receptor WYTD-containing repeat homology <YW19>  
F/1671-1714/Domain: LDL receptor WYTD-containing repeat homology <YW20>  
F/1715-1754/Domain: LDL receptor WYTD-containing repeat homology <YW21>  
F/1755-1797/Domain: LDL receptor WYTD-containing repeat homology <YW22>  
F/1798-1847/Domain: LDL receptor WYTD-containing repeat homology <YW23>  
F/1851-1887/Domain: EGF homology <EG8>  
F/1935-1977/Domain: LDL receptor WYTD-containing repeat homology <YW24>  
F/1978-2020/Domain: LDL receptor WYTD-containing repeat homology <YW25>  
F/2021-2064/Domain: LDL receptor WYTD-containing repeat homology <YW26>  
F/2065-2106/Domain: LDL receptor WYTD-containing repeat homology <YW27>  
F/2107-2152/Domain: LDL receptor WYTD-containing repeat homology <YW28>  
F/2160-2195/Domain: EGF homology <EG9>  
F/2200-2242/Domain: LDL receptor WYTD-containing repeat homology <YW29>  
F/2254-2295/Domain: LDL receptor WYTD-containing repeat homology <YW30>  
F/2345-2389/Domain: LDL receptor WYTD-containing repeat homology <YW31>  
F/2390-2430/Domain: LDL receptor WYTD-containing repeat homology <YW32>  
F/2431-2474/Domain: LDL receptor WYTD-containing repeat homology <YW33>  
F/2483-2518/Domain: EGF homology <EG10>  
F/2525-2562/Domain: LDL receptor ligand-binding repeat homology <LDLB>  
F/2567-2601/Domain: LDL receptor ligand-binding repeat homology <LDLC>  
F/2606-2640/Domain: LDL receptor ligand-binding repeat homology <LDLD>  
F/2653-2689/Domain: LDL receptor ligand-binding repeat homology <LDLE>  
F/2697-2731/Domain: LDL receptor ligand-binding repeat homology <LDLF>  
F/2735-2770/Domain: LDL receptor ligand-binding repeat homology <LDLG>  
F/2775-2813/Domain: LDL receptor ligand-binding repeat homology <LDLH>  
F/2819-2854/Domain: LDL receptor ligand-binding repeat homology <LDLI>  
F/2859-2898/Domain: LDL receptor ligand-binding repeat homology <LDLJ>  
F/2905-2940/Domain: LDL receptor ligand-binding repeat homology <LDLK>  
F/2945-2981/Domain: EGF homology <EG11>  
F/2987-3022/Domain: LDL receptor WYTD-containing repeat homology <YW34>  
F/3030-3069/Domain: LDL receptor WYTD-containing repeat homology <YW35>  
F/3070-3114/Domain: LDL receptor WYTD-containing repeat homology <YW36>

F/3115-3157/Domain: LDL receptor WYTD-containing repeat homology <YW36>  
F/3158-3201/Domain: LDL receptor WYTD-containing repeat homology <YW37>  
F/3202-3242/Domain: LDL receptor WYTD-containing repeat homology <YW38>  
F/3243-3285/Domain: LDL receptor WYTD-containing repeat homology <YW39>  
F/3285-3331/Domain: EGF homology <EG13>  
F/3335-3370/Domain: LDL receptor ligand-binding repeat homology <LDL4>  
F/3375-3409/Domain: LDL receptor ligand-binding repeat homology <LDL5>  
F/3414-3449/Domain: LDL receptor ligand-binding repeat homology <LDL6>  
F/3454-3490/Domain: LDL receptor ligand-binding repeat homology <LDL7>  
F/3495-3532/Domain: LDL receptor ligand-binding repeat homology <LDL8>  
F/3537-3571/Domain: LDL receptor ligand-binding repeat homology <LDL9>  
F/3576-3610/Domain: LDL receptor ligand-binding repeat homology <LDL10>  
F/3614-3648/Domain: LDL receptor ligand-binding repeat homology <LDL11>  
F/3655-3691/Domain: LDL receptor ligand-binding repeat homology <LDL12>  
F/3696-3732/Domain: LDL receptor ligand-binding repeat homology <LDL13>  
F/3742-3777/Domain: LDL receptor ligand-binding repeat homology <LDL14>  
F/3786-3823/Domain: EGF homology <EG14>  
F/3829-3861/Domain: EGF homology <EG15>  
F/3869-3912/Domain: LDL receptor WYTD-containing repeat homology <YW40>  
F/3913-3970/Domain: LDL receptor WYTD-containing repeat homology <YW41>  
F/3945-4545/Domain: alpha-2-macroglobulin receptor 85K chain #status predicted <85K>  
F/3945-4421/Domain: 85K chain extracellular #status predicted <EXT>  
F/3971-4013/Domain: LDL receptor WYTD-containing repeat homology <YW42>  
F/4014-4057/Domain: LDL receptor WYTD-containing repeat homology <YW43>  
F/4058-4100/Domain: LDL receptor WYTD-containing repeat homology <YW44>  
F/4101-4143/Domain: LDL receptor WYTD-containing repeat homology <YW45>  
F/4152-4183/Domain: EGF homology <EG16>  
F/4201-4232/Domain: EGF homology <EG17>  
F/4237-4268/Domain: EGF homology <EG18>  
F/4273-4304/Domain: EGF homology <EG19>  
F/4309-4340/Domain: EGF homology <EG20>  
F/4345-4375/Domain: EGF homology <EG21>  
F/4378-4409/Domain: EGF homology <EG22>  
F/4422-4445/Domain: transmembrane #status predicted <TM>  
F/4446-4545/Domain: intracellular #status predicted <INT>  
F/167,2999/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted  
F/2959/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted  
F/4076,4126,4279/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 96.0%; Score 486; DB 1; Length 4545;  
Best Local Similarity 95.3%; Pred. No. 3e-36; Indels 0; Gaps 0;  
Matches 82; Conservative 2; Mismatches 2

QY 1 KTCSPKPFACRQDQITCISKWRCGERDCPDGSDAEIQCPSKAQRCQPNHNCLETTEL 60  
26 KTCSPKPFACRQDQITCISKWRCGERDCPDGSDAEIQCPSKAQRCPPNHSCLETTEL 85

QY 61 CVPMSRLCNGVQDCWDSDEGPHORE 86  
86 CVPMSRLCNGVQDCWDSDEGAHORE 111

RESULT 3  
A3102  
Alpha-2-macroglobulin receptor precursor - chicken  
N/Alternate names: CD91; LDL receptor-related protein 1; low density lipoprotein recept  
C/Species: Gallus gallus (chicken)  
C/Date: 04-Sep-1998 #sequence\_revision 04-Sep-1998 #text\_change 22-Jun-1999  
C/Accession: A53102  
R.Nimf, J.; Stifani, S.; Bilous, P.T.; Schneider, W.J.  
J. Biol. Chem. 269, 212-219, 1994  
A/Title: The somatic cell-specific low density lipoprotein receptor-related protein of  
C/Reference number: A53102; MUID:94103212; PMID:7506255  
A/Accession: A53102  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-4543 <NIM>  
A/Cross-references: GS:X74904; NID:G438006; PIDN:CAA52870.1; PID:G438007  
C/Complex: The alpha-2-macroglobulin receptor complex consists of noncovalently-associat  
d protein.  
C/Superfamily: alpha-2-macroglobulin receptor; EGF homology; LDL receptor ligand-binding  
C/Keywords: beta-hydroxyasparagine; beta-hydroxyaspartic acid; calcium binding; glycopr  
F/1-17/Domain: signal sequence #status predicted <SIG>

F:18-3942/Domain: alpha-2-macroglobulin receptor S15K chain #status predicted <S15K>  
 F:18-3942.3943-4543/Product: alpha-2-macroglobulin receptor #status predicted <MAT>  
 F:28-66/Domain: LDL receptor ligand-binding repeat homology <LDL1>  
 F:74-110/Domain: LDL receptor ligand-binding repeat homology <LDL2>  
 F:117-150/Domain: EGF homology <EG1>  
 F:156-190/Domain: EGF homology <EG2>  
 F:200-241/Domain: LDL receptor WYTD-containing repeat homology <YW01>  
 F:242-283/Domain: LDL receptor WYTD-containing repeat homology <YW02>  
 F:294-336/Domain: LDL receptor WYTD-containing repeat homology <YW03>  
 F:337-380/Domain: LDL receptor WYTD-containing repeat homology <YW04>  
 F:381-422/Domain: LDL receptor WYTD-containing repeat homology <YW05>  
 F:423-470/Domain: LDL receptor WYTD-containing repeat homology <YW06>  
 F:480-521/Domain: EGF homology <EG3>  
 F:573-615/Domain: LDL receptor WYTD-containing repeat homology <YW07>  
 F:616-661/Domain: LDL receptor WYTD-containing repeat homology <YW08>  
 F:662-712/Domain: LDL receptor WYTD-containing repeat homology <YW09>  
 F:713-754/Domain: LDL receptor WYTD-containing repeat homology <YW10>  
 F:755-797/Domain: LDL receptor WYTD-containing repeat homology <YW11>  
 F:805-840/Domain: EGF homology <EG4>  
 F:852-888/Domain: LDL receptor ligand-binding repeat homology <LDL3>  
 F:893-929/Domain: LDL receptor ligand-binding repeat homology <LDL4>  
 F:934-969/Domain: LDL receptor ligand-binding repeat homology <LDL5>  
 F:974-1009/Domain: LDL receptor ligand-binding repeat homology <LDL6>  
 F:1013-1049/Domain: LDL receptor ligand-binding repeat homology <LDL7>  
 F:1060-1095/Domain: LDL receptor ligand-binding repeat homology <LDL8>  
 F:1102-1138/Domain: LDL receptor ligand-binding repeat homology <LDL9>  
 F:1143-1180/Domain: LDL receptor ligand-binding repeat homology <LDL10>  
 F:1183-1219/Domain: EGF homology <EG5>  
 F:1225-1259/Domain: EGF homology <EG6>  
 F:1267-1306/Domain: LDL receptor WYTD-containing repeat homology <YW12>  
 F:1307-1353/Domain: LDL receptor WYTD-containing repeat homology <YW13>  
 F:1354-1396/Domain: LDL receptor WYTD-containing repeat homology <YW14>  
 F:1447-1486/Domain: LDL receptor WYTD-containing repeat homology <YW15>  
 F:1487-1529/Domain: LDL receptor WYTD-containing repeat homology <YW16>  
 F:1538-1576/Domain: EGF homology <EG7>  
 F:1581-1624/Domain: LDL receptor WYTD-containing repeat homology <YW18>  
 F:1625-1667/Domain: LDL receptor WYTD-containing repeat homology <YW19>  
 F:1668-1711/Domain: LDL receptor WYTD-containing repeat homology <YW20>  
 F:1712-1751/Domain: LDL receptor WYTD-containing repeat homology <YW21>  
 F:1752-1794/Domain: LDL receptor WYTD-containing repeat homology <YW22>  
 F:1795-1842/Domain: LDL receptor WYTD-containing repeat homology <YW23>  
 F:1846-1882/Domain: EGF homology <EG8>  
 F:1930-1972/Domain: LDL receptor WYTD-containing repeat homology <YW24>  
 F:1973-2015/Domain: LDL receptor WYTD-containing repeat homology <YW25>  
 F:2016-2059/Domain: LDL receptor WYTD-containing repeat homology <YW26>  
 F:2060-2101/Domain: LDL receptor WYTD-containing repeat homology <YW27>  
 F:2102-2147/Domain: LDL receptor WYTD-containing repeat homology <YW28>  
 F:2155-2190/Domain: EGF homology <EG9>  
 F:2195-2237/Domain: LDL receptor WYTD-containing repeat homology <YW29>  
 F:2247-2288/Domain: LDL receptor WYTD-containing repeat homology <YW30>  
 F:2338-2382/Domain: LDL receptor WYTD-containing repeat homology <YW31>  
 F:2383-2423/Domain: LDL receptor WYTD-containing repeat homology <YW32>  
 F:2424-2467/Domain: LDL receptor WYTD-containing repeat homology <YW33>  
 F:2476-2511/Domain: EGF homology <EG10>  
 F:2518-2555/Domain: LDL receptor ligand-binding repeat homology <LDL11>  
 F:2560-2594/Domain: LDL receptor ligand-binding repeat homology <LDL12>  
 F:2599-2633/Domain: LDL receptor ligand-binding repeat homology <LDL13>  
 F:2646-2682/Domain: LDL receptor ligand-binding repeat homology <LDL14>  
 F:2690-2724/Domain: LDL receptor ligand-binding repeat homology <LDL15>  
 F:2732-2767/Domain: LDL receptor ligand-binding repeat homology <LDL16>  
 F:2772-2810/Domain: LDL receptor ligand-binding repeat homology <LDL17>  
 F:2816-2851/Domain: LDL receptor ligand-binding repeat homology <LDL18>  
 F:2856-2895/Domain: LDL receptor ligand-binding repeat homology <LDL19>  
 F:2902-2936/Domain: LDL receptor ligand-binding repeat homology <LDL20>  
 F:2941-2977/Domain: LDL receptor ligand-binding repeat homology <LDL21>  
 F:2983-3018/Domain: EGF homology <EG11>  
 F:3026-3065/Domain: LDL receptor WYTD-containing repeat homology <YW34>  
 F:3066-3110/Domain: LDL receptor WYTD-containing repeat homology <YW35>  
 F:3111-3153/Domain: LDL receptor WYTD-containing repeat homology <YW36>  
 F:3154-3197/Domain: LDL receptor WYTD-containing repeat homology <YW37>  
 F:3198-3238/Domain: LDL receptor WYTD-containing repeat homology <YW38>  
 F:3239-3281/Domain: LDL receptor WYTD-containing repeat homology <YW39>

F:3291-3327/Domain: EGF homology <EG13>  
 F:3331-3366/Domain: LDL receptor ligand-binding repeat homology <LDL1>  
 F:3371-3405/Domain: LDL receptor ligand-binding repeat homology <LDL2>  
 F:3410-3445/Domain: LDL receptor ligand-binding repeat homology <LDL3>  
 F:3450-3486/Domain: LDL receptor ligand-binding repeat homology <LDL4>  
 F:3491-3528/Domain: LDL receptor ligand-binding repeat homology <LDL5>  
 F:3533-3567/Domain: LDL receptor ligand-binding repeat homology <LDL6>  
 F:3572-3606/Domain: LDL receptor ligand-binding repeat homology <LDL7>  
 F:3610-3644/Domain: LDL receptor ligand-binding repeat homology <LDL8>  
 F:3651-3687/Domain: LDL receptor ligand-binding repeat homology <LDL9>  
 F:3692-3728/Domain: LDL receptor ligand-binding repeat homology <LDL10>  
 F:3738-3774/Domain: LDL receptor ligand-binding repeat homology <LDL11>  
 F:3783-3820/Domain: EGF homology <EG14>  
 F:3826-3858/Domain: EGF homology <EG15>  
 F:3866-3909/Domain: LDL receptor WYTD-containing repeat homology <YW40>  
 F:3910-3968/Domain: LDL receptor WYTD-containing repeat homology <YW41>  
 F:3943-4543/Domain: alpha-2-macroglobulin receptor 85K chain #status predicted <85K>  
 F:3943-4420/Domain: 85K chain extracellular #status predicted <EXT>  
 F:3959-4011/Domain: LDL receptor WYTD-containing repeat homology <YW42>  
 F:4012-4055/Domain: LDL receptor WYTD-containing repeat homology <YW43>  
 F:4056-4098/Domain: LDL receptor WYTD-containing repeat homology <YW44>  
 F:4099-4141/Domain: LDL receptor WYTD-containing repeat homology <YW45>  
 F:4150-4181/Domain: EGF homology <EG16>  
 F:4159-4230/Domain: EGF homology <EG17>  
 F:4235-4266/Domain: EGF homology <EG18>  
 F:4271-4302/Domain: EGF homology <EG19>  
 F:4307-4338/Domain: EGF homology <EG20>  
 F:4343-4373/Domain: EGF homology <EG21>  
 F:4376-4408/Domain: EGF homology <EG22>  
 F:4421-4443/Domain: transmembrane #status predicted <TM>  
 F:4444-4543/Domain: intracellular #status predicted <INT>  
 F:116,138,187,276,359,448,731,926,1048,1152,1153,1193,1216,1305,1509,1556,1573,1614,1643  
 3485,3659,3786,3837,3952,4074,4124,4178,4278/Binding site: carboxydrate (Asn) (covalent)  
 F:168,2395/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted  
 F:2955/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted

Query Match 90.7%; Score 459; DB 1; Length 4543;  
 Best Local Similarity 86.0%; Pred. No. 8 4e-34;  
 Matches 74; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Qy 1 KTCSPKQFACRQDITICISKWRCDGSRDPCDGSDEAPBICPQSKAORCPNEHNCGLTEL 60  
 Db 27 KTCSPKQFACKDQITICISKWRCDGSRDPCDGSDESPDICPQSKVSRCPNEHNCGLTEL 86  
 Qy 61 CVPMRLCNGVQDMGSDSGPHCRE 86  
 Db 87 CIEMSLCNGHDCFGSDSGPHCRE 112

RESULT 4  
 A47437  
 LDL-receptor-related protein - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text\_change 18-Aug-2000  
 C/Accession: A47437; S27801; T21547  
 R/Yochim, J.; Greenwald, I.  
 Proc. Natl. Acad. Sci. U.S.A. 90, 4572-4576, 1993  
 A/Title: A gene for a low density lipoprotein receptor-related protein in the nematode C.  
 A/Reference number: A47437; MUID:93281621; PMID:8506301  
 A/Accession: A47437  
 A/Molecule type: DNA  
 A/Residues: 1-4753 <YOC>  
 A/Cross-references: GB:M96150; NID:g156359; PIDN:AAA28105.1; PID:g156360  
 A/Note: nucleotide sequence not given; translation not complete in this paper  
 R/Yochim, J.; Greenwald, I.  
 submitted to the EMBL Data Library, July 1992  
 A/Description: A gene for an LDL receptor-related protein (LPR) in the nematode C.elegans  
 A/Reference number: S27801  
 A/Accession: S27801  
 A/Molecule type: DNA  
 A/Residues: 1-4753 <YOC>  
 A/Cross-references: EMBL:M96150; NID:g156359; PIDN:AAA28105.1; PID:g156360  
 R/Wilkinson, J.

submitted to the EMBL Data Library, June 1996

A;Reference number: Z19439  
A;Accession: T21547  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-4753 <WIL>  
A;Cross-references: EMBL:Z73907; PIDN:CAA98124.1; GSPDB:GN00015; CESP:F29D11.1  
A;Experimental source: clone F29D11  
C;Genetics:  
A;Gene: LPR  
A;Map position: 1  
A;Introns: 31/1; 88/1; 132/1; 172/3; 219/1; 298/1; 463/2; 526/2; 585/3; 780/2; 874/2; 974/2; 151/1  
C;Superfamily: alpha-2-macroglobulin receptor; EGF homology; LDL receptor ligand-binding  
C;Keywords: tandem repeat; transmembrane protein  
F;33-87/Domain: LDL receptor ligand-binding repeat homology <LDL1>  
F;92-131/Domain: LDL receptor ligand-binding repeat homology <LDL2>  
F;138-175/Domain: LDL receptor ligand-binding repeat homology <LDL3>  
F;182-218/Domain: LDL receptor ligand-binding repeat homology <LDL4>  
F;223-257/Domain: LDL receptor ligand-binding repeat homology <LDL5>  
F;262-297/Domain: LDL receptor ligand-binding repeat homology <LDL6>  
F;302-336/Domain: EGF homology <EGF1>  
F;1054-1095/Domain: LDL receptor ligand-binding repeat homology <LDL7>  
F;1101-1138/Domain: LDL receptor ligand-binding repeat homology <LDL8>  
F;1146-1182/Domain: LDL receptor ligand-binding repeat homology <LDL9>  
F;1187-1223/Domain: LDL receptor ligand-binding repeat homology <LDL10>  
F;1228-1263/Domain: LDL receptor ligand-binding repeat homology <LDL11>  
F;1270-1307/Domain: LDL receptor ligand-binding repeat homology <LDL12>  
F;1313-1350/Domain: LDL receptor ligand-binding repeat homology <LDL13>  
F;1359-1396/Domain: LDL receptor ligand-binding repeat homology <LDL14>  
F;1441-1475/Domain: EGF homology <EGF>  
F;1511-1554/Domain: LDL receptor WYTD-containing repeat homology <YW33>  
F;2792-2829/Domain: LDL receptor ligand-binding repeat homology <LDL15>  
F;2834-2868/Domain: LDL receptor ligand-binding repeat homology <LDL16>  
F;2874-2912/Domain: LDL receptor ligand-binding repeat homology <LDL17>  
F;2919-2956/Domain: LDL receptor ligand-binding repeat homology <LDL18>  
F;2961-2997/Domain: LDL receptor ligand-binding repeat homology <LDL19>  
F;3006-3044/Domain: LDL receptor ligand-binding repeat homology <LDL20>  
F;3049-3093/Domain: LDL receptor ligand-binding repeat homology <LDL21>  
F;3100-3135/Domain: LDL receptor ligand-binding repeat homology <LDL22>  
F;3140-3174/Domain: LDL receptor ligand-binding repeat homology <LDL23>  
F;3187-3222/Domain: LDL receptor ligand-binding repeat homology <LDL24>  
F;3586-3623/Domain: EGF homology <EGF1>  
F;3627-3666/Domain: LDL receptor ligand-binding repeat homology <LDL25>  
F;3671-3705/Domain: LDL receptor ligand-binding repeat homology <LDL26>  
F;3709-3746/Domain: LDL receptor ligand-binding repeat homology <LDL27>  
F;3753-3788/Domain: LDL receptor ligand-binding repeat homology <LDL28>  
F;3793-3830/Domain: LDL receptor ligand-binding repeat homology <LDL29>  
F;3833-3871/Domain: LDL receptor ligand-binding repeat homology <LDL30>  
F;3878-3912/Domain: LDL receptor ligand-binding repeat homology <LDL31>  
F;3917-3951/Domain: LDL receptor ligand-binding repeat homology <LDL32>  
F;3959-3995/Domain: LDL receptor ligand-binding repeat homology <LDL33>  
F;4000-4040/Domain: LDL receptor ligand-binding repeat homology <LDL34>  
F;4049-4083/Domain: LDL receptor ligand-binding repeat homology <LDL35>  
F;4092-4130/Domain: EGF homology <EGF2>  
F;4343-4386/Domain: LDL receptor WYTD-containing repeat homology <YW38>  
  
Query Match 40.4%; Score 204.5; DB 1; Length 4753;  
Best Local Similarity 36.0%; Pred. No. 1.1e-10;  
Matches 41; Conservative 10; Mismatches 26; Indels 37; Gaps 3;  
  
Qy 3 CSPKQFACDQITCISKGWRCDEBPCDGSDEAPETCP-----QSKAQRCPNHH 53  
Db 182 CFQYQFACDQITCISKGWRCDEBPCDGSDE--PDCFEKKTANEFCNKRCPKRF 240  
Qy 54 -----NCLGTCLCPMSRLCNGVQDCMDGSDS 80  
Db 241 RCDYDDGSDNDECEGYRCRPGKWNCPGTGHCIDQLKCDGSKDCADGAD 294  
  
RESULT 5  
T42737  
gp330 protein precursor - rat

N;Alternate names: megalin  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 04-Mar-2000  
C;Accession: T42737  
R;Saito, A.; Pietromonaco, S.; Loo, A.K.C.; Farquhar, M.G.  
Proc. Natl. Acad. Sci. U.S.A. 91, 9725-9729, 1994  
A;Title: Complete cloning and sequencing of rat gp330/megalin, a distinctive member of  
A;Reference number: A58173; MUID:95024033; PMID:7937880  
A;Accession: T42737  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-4660 <BAI>  
A;Cross-references: EMBL:L34049; MID:9561852; PID:AAAS1369.1  
A;Experimental source: strain Sprague-Dawley; kidney  
C;Superfamily: alpha-2-macroglobulin receptor; EGF homology; LDL receptor ligand-binding  
F;1-25/Domain: signal sequence #status Predicted <SIG>  
F;26-4660/Product: gp330 protein #status Predicted <MAT>  
  
Query Match 39.9%; Score 202; DB 2; Length 4660;  
Best Local Similarity 45.3%; Pred. No. 1.8e-10;  
Matches 39; Conservative 11; Mismatches 32; Indels 4; Gaps 3;  
  
Qy 1 KTCSPKQFACDQITCISKGWRCDEBPCDGSDEAPETCPQSKAQRCPNHHNCLGTCL 60  
Db 1270 KTCSPFTFLC--DNGNCIYKAWICDGDNDRCRMSDE--KDCP-TQFFHCPSTQWCPGYST 1325  
Qy 61 CVPMSRLCNGVQDCMDGSDGEPHCRC 86  
Db 1326 CINLSALCDGVDFCNGTDESPLCNQ 1351  
  
RESULT 6  
QXLLI2  
LDL receptor 2 precursor - African clawed frog  
C;Species: Xenopus laevis (African clawed frog)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 18-Sep-1998  
C;Accession: B40388  
R;Menta, K.D.; Chen, W.J.; Goldstein, J.L.; Brown, M.S.  
J. Biol. Chem. 266, 10406-10414, 1991  
A;Title: The low density lipoprotein receptor in Xenopus laevis. Five domains that resea  
A;Reference number: A40388; MUID:91244815; PMID:1709931  
A;Accession: B40388  
A;Molecule type: mRNA  
A;Residues: 1-909 <MEH>  
A;Cross-references: GB:M62979  
A;Note: The authors translated the codon AAA for residue 630 as Asn  
nd complexes must first cluster into clathrin-coated pits.  
C;Superfamily: LDL receptor; EGF homology; LDL receptor ligand-binding repeat homology;  
C;Keywords: cholesterol; coated pits; duplication; endocytosis; glycoprotein; LDL; lipi  
F;1-21/Domain: signal sequence #status Predicted <SIG>  
F;22-909/Product: LDL receptor #status Predicted <EXT>  
F;22-836/Domain: extracellular #status Predicted <EXT>  
F;27-63/Domain: LDL receptor ligand-binding repeat homology <LDL1>  
F;68-104/Domain: LDL receptor ligand-binding repeat homology <LDL2>  
F;109-143/Domain: LDL receptor ligand-binding repeat homology <LDL3>  
F;148-183/Domain: LDL receptor ligand-binding repeat homology <LDL4>  
F;195-229/Domain: LDL receptor ligand-binding repeat homology <LDL5>  
F;234-268/Domain: LDL receptor ligand-binding repeat homology <LDL6>  
F;274-311/Domain: LDL receptor ligand-binding repeat homology <LDL7>  
F;316-350/Domain: EGF homology <EG1>  
F;356-390/Domain: EGF homology <EG2>  
F;397-436/Domain: LDL receptor WYTD-containing repeat homology <YW1>  
F;437-483/Domain: LDL receptor WYTD-containing repeat homology <YW2>  
F;484-526/Domain: LDL receptor WYTD-containing repeat homology <YW3>  
F;527-570/Domain: LDL receptor WYTD-containing repeat homology <YW4>  
F;571-613/Domain: LDL receptor WYTD-containing repeat homology <YW5>  
F;614-656/Domain: LDL receptor WYTD-containing repeat homology <YW6>  
F;665-709/Domain: EGF homology <EG3>  
F;717-813/Region: clustered O-linked oligosaccharides  
F;837-858/Domain: transmembrane #status Predicted <TMM>  
F;859-909/Domain: intracellular #status Predicted <INT>  
F;873-877/Region: coated-pit mediated internalization signal



## RESULT 10

ORXLL1  
LDL receptor 1 precursor - African clawed frog  
C/Species: Xenopus laevis (African clawed frog)  
C/Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 18-Sep-1998  
C/Accession: A40388  
R/Mehta, K.D.; Chen, W.J.; Goldstein, J.L.; Brown, M.S.  
J. Biol. Chem. 266, 10406-10414, 1991  
A/Title: The low density lipoprotein receptor in Xenopus laevis. Five domains that resemble the structure of the hamster low density lipoprotein receptor gene.  
A/Reference number: A40388; MUID:91244815; PMID:1709931  
A/Accession: A40388  
A/Molecule type: mRNA  
A/Residues: 1-909 <MEH>  
A/Cross-references: GB:M62976  
C/Comment: This transmembrane glycoprotein binds LDL, the major cholesterol-carrying lipoprotein complex must first cluster into clathrin-coated pits.  
C/Superfamily: LDL receptor; EGF homology; LDL receptor ligand-binding repeat homology; LDL receptor; cholesterol; coated pits; duplication; endocytosis; Glycoprotein; LDL; lipid  
C/Keywords: cholesterol; coated pits; duplication; endocytosis; Glycoprotein; LDL; lipid  
F/1-21/Domain: signal sequence #status predicted <SIG>  
F/22-909/Product: LDL receptor #status predicted <MAT>  
F/22-836/Domain: extracellular #status predicted <EXT>  
F/27-63/Domain: LDL receptor ligand-binding repeat homology <LDL1>  
F/68-104/Domain: LDL receptor ligand-binding repeat homology <LDL2>  
F/109-143/Domain: LDL receptor ligand-binding repeat homology <LDL3>  
F/148-183/Domain: LDL receptor ligand-binding repeat homology <LDL4>  
F/195-229/Domain: LDL receptor ligand-binding repeat homology <LDL5>  
F/234-268/Domain: LDL receptor ligand-binding repeat homology <LDL6>  
F/274-311/Domain: LDL receptor ligand-binding repeat homology <LDL7>  
F/316-350/Domain: EGF homology <EG1>  
F/356-390/Domain: EGF homology <EG2>  
F/397-436/Domain: LDL receptor WYTD-containing repeat homology <YW1>  
F/437-483/Domain: LDL receptor WYTD-containing repeat homology <YW2>  
F/484-526/Domain: LDL receptor WYTD-containing repeat homology <YW3>  
F/527-570/Domain: LDL receptor WYTD-containing repeat homology <YW4>  
F/571-613/Domain: LDL receptor WYTD-containing repeat homology <YW5>  
F/614-656/Domain: LDL receptor WYTD-containing repeat homology <YW6>  
F/665-709/Domain: EGF homology <EG3>  
F/717-813/Region: clustered O-linked oligosaccharides  
F/837-858/Domain: transmembrane #status predicted <TM>  
F/859-909/Domain: intracellular #status predicted <INT>  
F/873-877/Region: coated-pit mediated internalization signal  
F/896-898/Region: basolateral targeting signal  
F/97-270/459/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F/316-327,323-336,338-350,356-366,362-375,377-390,665-679,675-694,696-709/Disulfide bond

Query Match 38.0%; Score 192.5; DB 1; Length 909;  
Best Local Similarity 45.7%; Pred. No. 3.7e-10;  
Matches 42; Conservative 9; Mismatches 30; Indels 11; Gaps 4;

Qy 2 TCSPKQFACRQITCISKWRCGDERCPDGSDEAPEC---PQSKAQRCPNEHNCUG 57  
Db 147 TCNPFQCKDKGICIPKMWACDPPDCBDSDE--EHCEGRPIKTDKPCSPLEPHC-G 203

Qy 58 TELCPVMSRLCNGVQDCMGDSDE---GPHCR 85  
Db 204 SGECIHSWKCDGGFDCKDKSDEKCVKPCR 235

submitted to the EMBL Data Library, December 1988  
A/Reference number: S05076  
A/Accession: S05076  
A/Molecule type: mRNA  
A/Residues: 1-187,'D',189-879 <LE2>  
A/Cross-references: EMBL:X13722; NID:956569; PIDN:CAA32001.1; PID:956570  
R/Bishop, R.W.  
J. Lipid Res. 33, 549-557, 1992  
A/Title: Structure of the hamster low density lipoprotein receptor gene.  
A/Reference number: A48908; MUID:92407466; PMID:1527478  
A/Accession: B48908  
A/Molecule type: DNA  
A/Residues: 1-7 <BIS>  
C/Comment: This transmembrane glycoprotein binds LDL, the major cholesterol-carrying lipoprotein complex must first cluster into clathrin-coated pits.  
C/Superfamily: LDL receptor; EGF homology; LDL receptor ligand-binding repeat homology; LDL receptor; cholesterol; coated pits; duplication; endocytosis; Glycoprotein; LDL; lipid  
C/Keywords: cholesterol; coated pits; duplication; endocytosis; Glycoprotein; LDL; lipid  
F/1-21/Domain: signal sequence #status predicted <SIG>  
F/22-879/Product: LDL receptor #status predicted <MAT>  
F/22-807/Domain: extracellular #status predicted <EXT>  
F/27-63/Domain: LDL receptor ligand-binding repeat homology <LDL1>  
F/68-104/Domain: LDL receptor ligand-binding repeat homology <LDL2>  
F/109-143/Domain: LDL receptor ligand-binding repeat homology <LDL3>  
F/148-184/Domain: LDL receptor ligand-binding repeat homology <LDL4>  
F/198-232/Domain: LDL receptor ligand-binding repeat homology <LDL5>  
F/237-271/Domain: LDL receptor ligand-binding repeat homology <LDL6>  
F/277-314/Domain: LDL receptor ligand-binding repeat homology <LDL7>  
F/319-353/Domain: EGF homology <EG1>  
F/359-393/Domain: EGF homology <EG2>  
F/400-439/Domain: LDL receptor WYTD-containing repeat homology <YW1>  
F/440-485/Domain: LDL receptor WYTD-containing repeat homology <YW2>  
F/486-528/Domain: LDL receptor WYTD-containing repeat homology <YW3>  
F/529-572/Domain: LDL receptor WYTD-containing repeat homology <YW4>  
F/573-615/Domain: LDL receptor WYTD-containing repeat homology <YW5>  
F/616-658/Domain: LDL receptor WYTD-containing repeat homology <YW6>  
F/667-711/Domain: EGF homology <EG3>  
F/717-788/Region: clustered O-linked oligosaccharides  
F/808-829/Domain: transmembrane #status predicted <TM>  
F/830-879/Domain: intracellular #status predicted <INT>  
F/844-848/Region: coated-pit mediated internalization signal  
F/856-868/Region: basolateral targeting signal  
F/97-156,273,657/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F/319-330,326-339,341-353,359-369,365-378,380-393,667-681,677-696,698-711/Disulfide bond  
F/717,720,721,724,726,733,748,757,766,768,775,780,785/Binding site: carbohydrate (T)  
F/725,738,738,746,755,764,767,769,779,783,787,788/Binding site: carbohydrate (Ser) (cov

Query Match 37.8%; Score 191.5; DB 1; Length 879;  
Best Local Similarity 46.2%; Pred. No. 4.5e-10;  
Matches 37; Conservative 10; Mismatches 28; Indels 5; Gaps 3;

Qy 2 TCSPKQFACRQITCISKWRCGDERCPDGSDEAPECQSKAQRCPNEHNCUG-TEL 60  
Db 26 SCGNFEFCRDG-KCIIVSKWVCDGSRCPDGSDESPETC---MSVTCRSGFSCGGRVSR 81

Qy 61 CVPMSRLCNGVQDCMGDSDE 80  
Db 82 CIPDSWRCDGRTDCENGSD 101

## RESULT 12

T29149  
LDL receptor precursor - rat  
C/Species: Rattus norvegicus (Norway rat)  
C/Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 22-Jun-1999  
C/Accession: S03430; S03076; B48908  
R/Lee, L.Y.; Mchler, W.A.; Schafer, B.L.; Freudenberger, J.S.; Byrne-Connolly, N.; Eager  
Nucleic Acids Res. 17, 1259-1260, 1989  
A/Title: Nucleotide sequence of the rat low density lipoprotein receptor cDNA.  
A/Reference number: S03430; MUID:89160263; PMID:2922268  
A/Accession: S03430  
A/Molecule type: mRNA  
A/Residues: 1-879 <LEE>  
A/Cross-references: EMBL:X13722  
R/Lee, L.



A;Cross-references: EMBL:U97017; PIDN:AA52363.1; GSPDB:GN00019; CESP:F47B3.8  
A;Experimental source: strain Bristol N2; clone F47B3  
C;Genetics:  
A;Gene: CESP:F47B3.8  
A;Map position: 1  
A;Introns: 29/1; 78/1; 124/1; 216/1; 256/1; 1222/3; 2036/3; 2127/2; 2478/3; 2522/3; 2589/3  
C;Superfamily: LDL receptor ligand-binding repeat homology  
F;44-77/Domain: LDL receptor ligand-binding repeat homology <LDL1>  
F;81-124/Domain: LDL receptor ligand-binding repeat homology <LDL2>  
F;130-166/Domain: LDL receptor ligand-binding repeat homology <LDL3>  
F;925-964/Domain: LDL receptor ligand-binding repeat homology <LDL4>  
F;969-1005/Domain: LDL receptor ligand-binding repeat homology <LDL5>  
F;1013-1047/Domain: LDL receptor ligand-binding repeat homology <LDL6>  
F;1053-1087/Domain: LDL receptor ligand-binding repeat homology <LDL7>  
F;1093-1132/Domain: LDL receptor ligand-binding repeat homology <LDL8>  
F;1140-1175/Domain: LDL receptor ligand-binding repeat homology <LDL9>  
F;1180-1216/Domain: LDL receptor ligand-binding repeat homology <LDL10>  
F;1228-1265/Domain: LDL receptor ligand-binding repeat homology <LDL11>  
  
Query Match 37.8%; Score 191.5; DB 2; Length 2643;  
Best Local Similarity 46.4%; Pred. No. 18-09;  
Matches 39; Conservative 13; Mismatches 25; Indels 7; Gaps 5;  
  
QY 3 CSPKQACRQDITCISKWRCGERDCPDGSEAPICPSKAQRCQPNHCLGTELCV 62  
Db 1013 CSVSQFQC-SKTKCIKSKRCNGVQCDNGADE--EDCFRSL--CDPDEFRC-GTGLCI 1066  
  
QY 63 FMSRLCNGVQDCMDGSGPHCRE 86  
Db 1067 KQSQVCDGKMQCLDLGLDE-EHCNE 1089  
  
RESULT 13  
QRMSLD  
LDL receptor precursor - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 31-Dec-1993 #sequence\_revision 14-Aug-1998 #text\_change 22-Jun-1999  
A;Accession: J48623; JN0461  
R;Polvino, W.J.; Dichek, D.A.; Mason, J.; Anderson, W.F.  
Somat. Cell Mol. Genet. 18, 443-450, 1992  
A;Title: Molecular cloning and nucleotide sequence of cDNA encoding a functional murine  
A;Reference number: I48623; MUID:93117934; PMID:1475710  
A;Accession: I48623  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-862 <RES>  
A;Cross-references: EMBL:X64414; NID:G296201; PIDN:CAA45759.1; PID:G296202  
R;Hoffer, M.J.V.; van Eck, M.W.; Petrij, F.; van der Zee, A.; de Wit, E.; Meijer, D.; GZ  
Biochem. Biophys. Res. Commun. 191, 880-886, 1993  
A;Title: The mouse low density lipoprotein receptor gene: cDNA sequence and exon-intron  
A;Reference number: JN0461; MUID:93221531; PMID:8466528  
A;Accession: JN0461  
A;Molecule type: mRNA  
A;Residues: 1-22 'V', 24-26 'G', 28-60 'K', 62-143 'P', 145-155 'K', 157-177 'H', 179-185 'AE'  
A;Cross-references: GB:Z19521; NID:G296378; PIDN:CAA79581.1; PID:G296379  
A;Experimental source: strain BAUB/C, adult male liver  
C;Comment: This transmembrane glycoprotein binds LDL, the major cholesterol-carrying lip  
id complexes must first cluster into clathrin-coated pits.  
C;Superfamily: LDL receptor; EGF homology; LDL receptor ligand-binding repeat homology;  
C;Keywords: cholesterol; coated pits; duplication; endocytosis; glycoprotein; LDL; lipid  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;22-862/Product: LDL receptor #status predicted <SIG>  
F;22-790/Domain: extracellular #status predicted <EXT>  
F;27-63/Domain: LDL receptor ligand-binding repeat homology <LDL1>  
F;68-104/Domain: LDL receptor ligand-binding repeat homology <LDL2>  
F;109-143/Domain: LDL receptor ligand-binding repeat homology <LDL3>  
F;148-184/Domain: LDL receptor ligand-binding repeat homology <LDL4>  
F;198-232/Domain: LDL receptor ligand-binding repeat homology <LDL5>  
F;237-271/Domain: LDL receptor ligand-binding repeat homology <LDL6>  
F;277-314/Domain: LDL receptor ligand-binding repeat homology <LDL7>  
F;319-353/Domain: EGF homology <EG1>  
F;359-393/Domain: EGF homology <EG2>  
F;400-439/Domain: LDL receptor YWTD-containing repeat homology <YW1>

F;440-485/Domain: LDL receptor YWTD-containing repeat homology <YW2>  
F;486-528/Domain: LDL receptor YWTD-containing repeat homology <YW3>  
F;529-572/Domain: LDL receptor YWTD-containing repeat homology <YW4>  
F;573-615/Domain: LDL receptor YWTD-containing repeat homology <YW5>  
F;616-658/Domain: LDL receptor YWTD-containing repeat homology <YW6>  
F;667-712/Domain: EGF homology <EG3>  
F;722-770/Region: clustered O-linked oligosaccharides  
F;791-812/Domain: transmembrane #status predicted <TM>  
F;813-862/Domain: intracellular #status predicted <INT>  
F;827-831/Region: coated-pit mediated internalization signal  
F;839-851/Region: basolateral targeting signal  
F;97-273,422/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;319-330,326-339,341-353,359-369,365-378,380-393,667-682,678-697,699-712/Disulfide bond  
  
Query Match 37.6%; Score 190.5; DB 1; Length 862;  
Best Local Similarity 46.2%; Pred. No. 5.4e-10;  
Matches 37; Conservative 11; Mismatches 27; Indels 5; Gaps 3;  
  
QY 2 TCSPKQACRQDITCISKWRCGERDCPDGSEAPICPSKAQRCQPNHCLG-TEL 60  
Db 26 SCSSNEFCQKRDG-KCTASKWVCDGSPCPDGSDEPETC---MSVTCQSNQFSGGRVSR 81  
  
QY 61 CVPMSRLCNGVQDCMDGSGDE 80  
Db 82 CIPDSWRCQVDCENDSDE 101  
  
RESULT 14  
QRMULD  
LDL receptor precursor - Chinese hamster  
C;Species: Cricetulus griseus (Chinese hamster)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 22-Jun-1999  
A;Accession: A48908; A24426  
R;Bishop, R.W.  
J. Lipid Res. 33, 549-557, 1992  
A;Title: Structure of the hamster low density lipoprotein receptor gene.  
A;Reference number: A48908; MUID:92407468; PMID:1527478  
A;Accession: A48908  
A;Molecule type: DNA  
A;Residues: 1-854 <BIS>  
A;Cross-references: GB:M4387; NID:G191131; PIDN:AA51449.1; PID:G191132  
A;Experimental source: UT-1 ovarian cell genomic library  
A;Note: sequence extracted from NCBI backbone (NCBIN:113276, NCBIIP:113277)  
R;Sege, R.D.; Kozarsky, K.F.; Krieger, M.  
Mol. Cell. Biol. 6, 3268-3277, 1986  
A;Title: Characterization of a family of gamma-ray-induced CHO mutants demonstrates that  
A;Reference number: A24426; MUID:87064645; PMID:3785227  
A;Accession: A24426  
A;Molecule type: DNA  
A;Residues: 570-615 <SEG>  
C;Genetics:  
A;Gene: ldlA  
A;Introns: 23/1; 64/1; 105/1; 233/1; 274/1; 315/1; 355/1; 397/1; 454/2; 529/2; 569/1; 61  
A;Superfamily: LDL receptor; EGF homology; LDL receptor ligand-binding repeat homology;  
C;Keywords: cholesterol; coated pits; duplication; endocytosis; glycoprotein; LDL; lipid  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;22-854/Product: LDL receptor #status predicted <EXT>  
F;27-63/Domain: LDL receptor ligand-binding repeat homology <LDL1>  
F;68-104/Domain: LDL receptor ligand-binding repeat homology <LDL2>  
F;109-143/Domain: LDL receptor ligand-binding repeat homology <LDL3>  
F;148-184/Domain: LDL receptor ligand-binding repeat homology <LDL4>  
F;198-232/Domain: LDL receptor ligand-binding repeat homology <LDL5>  
F;237-271/Domain: LDL receptor ligand-binding repeat homology <LDL6>  
F;277-314/Domain: LDL receptor ligand-binding repeat homology <LDL7>  
F;319-353/Domain: EGF homology <EG1>  
F;359-393/Domain: EGF homology <EG2>  
F;400-439/Domain: LDL receptor YWTD-containing repeat homology <YW1>  
F;440-485/Domain: LDL receptor YWTD-containing repeat homology <YW2>  
F;486-528/Domain: LDL receptor YWTD-containing repeat homology <YW3>  
F;529-572/Domain: LDL receptor YWTD-containing repeat homology <YW4>  
F;573-615/Domain: LDL receptor YWTD-containing repeat homology <YW5>  
F;616-658/Domain: LDL receptor YWTD-containing repeat homology <YW6>



Search completed: September 17, 2004, 11:18:37  
Job time : 40 secs

Query Match 37.4%; Score 189; DB 1; Length 854;  
Best Local Similarity 42.0%; Pred. No. 7.4e-10;  
Matches 37; Conservative 15; Mismatches 22; Indels 14; Gaps 5;

RESULT 15  
19764  
Prothetical protein T2L3.3 - *Caenorhabditis elegans*  
Species: *Caenorhabditis elegans*  
Date: 15-Oct-1999 #sequence revision 15-Oct-1999 #text change 18-Aug-2000

51-85/Domain: LDL receptor ligand-binding repeat homology <LDL1>  
90-132/Domain: LDL receptor ligand-binding repeat homology <LDL3>  
149-185/Domain: LDL receptor ligand-binding repeat homology <LDL4>  
190-225/Domain: LDL receptor ligand-binding repeat homology <LDL5>  
229-272/Domain: LDL receptor ligand-binding repeat homology <LDL6>  
285-316/Domain: LDL receptor ligand-binding repeat homology <LDL7>  
323-359/Domain: LDL receptor ligand-binding repeat homology <LDL8>  
369-405/Domain: LDL receptor ligand-binding repeat homology <LDL9>  
481-856/Domain: LDL receptor ligand-binding repeat homology <LDL10>  
861-903/Domain: LDL receptor ligand-binding repeat homology <LDL11>  
908-943/Domain: LDL receptor ligand-binding repeat homology <LDL12>  
948-987/Domain: LDL receptor ligand-binding repeat homology <LDL13>  
993-1027/Domain: LDL receptor ligand-binding repeat homology <LDL14>  
1031-1063/Domain: LDL receptor ligand-binding repeat homology <LDL15>  
1112-1165/Domain: LDL receptor ligand-binding repeat homology <LDL16>  
1172-1209/Domain: LDL receptor ligand-binding repeat homology <LDL17>  
1213-1246/Domain: LDL receptor ligand-binding repeat homology <LDL18>

```

Query Match      37.4%; Score 189; DB 2; Length 2180;
Best Local Similarity 42.0%; Pred. No. 1.5e-09;
Matches 34; Conservative 15; Mismatches 28; Indels 6; Gaps 2;

1 KTCSPK-QFACRDQITCISKWRCDGERDCPDGSDAEFICPQSKAQRCPQNEHNCIGTE 59
  :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
147 RSCDAKTOFACLATRTCPKHWQCDQDPDCADGDE-----KNCDAKCTGFSEFSCSSKK 201
  :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```

60  LCVPMRLCNGVQDCMDGSDE 80
    ||: : ||: : ||: ||
202 KCIPLEQKCDGRDPCDGEDE 222

```

Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2004, 11:02:49 ; Search time 24 Seconds  
(without alignments)  
186.585 Million cell updates/sec

Title: US-09-625-137-21

Perfect score: 506

Sequence: 1 KTCSPKQFACRDQITCSKG.....LCNGVQDCMDGSGDEGPHCRE 86

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.1\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	506	100.0	4544	1 LRPI_HUMAN	Q07954 homo sapien
2	459	90.7	4543	1 LRPI_CHICK	P98157 gallus gall
3	204.5	40.4	4753	1 LRP_CAEEL	Q04833 caenorhabdi
4	202	39.9	4660	1 LRP2_RAT	P98158 rattus norv
5	197.5	39.0	1984	1 YL_DROME	P98163 drosophila
6	192.5	38.0	509	1 LDLI_XENLA	Q99087 xenopus lae
7	191.5	37.8	879	1 LDLR_RAT	P35952 rattus norv
8	189	37.4	854	1 LDLR_CRIGR	P35950 cricetulus
9	187.5	37.1	2214	1 SORL_HUMAN	Q92673 h sortilin
10	185	36.6	3375	1 UNS2_CAEEL	Q06561 caenorhabdi
11	184.5	36.5	873	1 LDVR_MOUSE	P98156 mus musculu
12	184.5	36.5	2213	1 SORL_RABIT	Q95209 o sortilin
13	183	36.2	4555	1 LRP2_HUMAN	P98164 homo sapien
14	182.5	36.1	837	1 LDVR_RABIT	P20063 oryctolagus
15	182.5	36.1	863	1 LDVR_CHICK	P98165 gallus gall
16	182.5	36.1	873	1 LDVR_RABIT	P35953 oryctolagus
17	182.5	36.1	1592	1 SORL_CHICK	Q98930 g sortilin
18	181.5	35.9	873	1 LDVR_RAT	P98166 rattus norv
19	181.5	35.9	892	1 LDLI2_XENLA	Q99088 xenopus lae
20	180.5	35.7	873	1 LDVR_HUMAN	P98155 homo sapien
21	180	35.6	864	1 LDLR_MOUSE	P35951 mus musculu
22	179.5	35.5	2215	1 SORL_MOUSE	O88307 m sortilin
23	176	34.8	860	1 LDLR_HUMAN	P01130 homo sapien
24	174	34.4	4391	1 PGBM_HUMAN	P98160 homo sapien
25	165	32.6	3707	1 PGBM_MOUSE	Q05793 mus musculu
26	159.5	31.5	2616	1 NDL_DROME	P98159 drosophila
27	156.5	30.9	1113	1 COR1_MOUSE	Q92319 mus musculu
28	151.5	29.9	855	1 STI4_HUMAN	Q9Y5Y6 homo sapien
29	151	29.8	867	1 SSPO_BOVIN	P98167 bos taurus
30	141	27.9	855	1 STI4_MOUSE	P56677 mus musculu
31	140.5	27.8	1042	1 COR1_HUMAN	Q9Y5G5 homo sapien
32	139.5	27.6	1115	1 GPCR_LYMTST	P46023 lymnaea sca
33	137	27.1	3767	1 MUA3_CAEEL	P34576 caenorhabdi

ALIGNMENTS

RESULT 1

LRPI\_HUMAN

ID	LRPI_HUMAN	STANDARD;	PRT;	4544 AA.
AC	Q07954;			
DT	01-OCT-1996	(Rel. 34, Created)		
DT	01-OCT-1996	(Rel. 34, Last sequence update)		
DT	15-MAR-2004	(Rel. 43, Last annotation update)		
DE	Low-density lipoprotein receptor-related protein 1 precursor (LRP)			
DE	(Alpha-2-macroglobulin receptor) (A2MR) (Apolipoprotein E receptor)			
DE	(APOER) (CD91).			
GN	LRP1 OR A2MR.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
EX	MEDLINE=93210795; PubMed=3266596;			
RA	Herz J., Hamann U., Rognie S., Myklebost O., Gausepohl H.,			
RA	Stanley K.K.;			
RT	"Surface location and high affinity for calcium of a 500-kd liver			
RT	membrane protein closely related to the LDL-receptor suggest a			
RT	physiological role as lipoprotein receptor.";			
RL	EMBO J. 7:4119-4127(1988).			
RP	SEQUENCE FROM N.A.			
EX	MEDLINE=95203893; PubMed=7534747;			
RA	Van Leuven F., Stas L., Hilliker C., Lorent K., Umans L., Serneels L.,			
RA	Overbergh L., Torrekens S., Moechars D., De Strooper B.,			
RA	Van den Bergh H.;			
RT	"Structure of the gene (LRP1) coding for the human alpha			
RT	2-macroglobulin receptor lipoprotein receptor-related protein.";			
RL	Genomics 24:78-89(1994).			
RP	SEQUENCE FROM N.A.			
EX	MEDLINE=9900832; PubMed=9782078;			
RA	Van Leuven F., Stas L., Thiry E., Nelissen B., Miyake Y.;			
RT	"Strategy to sequence the 89 exons of the human LRP1 gene coding for			
RT	the lipoprotein receptor related protein: identification of one			
RT	expressed mutation among 48 polymorphisms.";			
RL	Genomics 52:138-144(1998).			
RP	PROCESSING.			
EX	MEDLINE=90269210; PubMed=2112085;			
RA	Herz J., Kowal R.C., Goldstein J.L., Brown M.S.;			
RT	"Proteolytic processing of the 600 kd low density lipoprotein			
RT	receptor-related protein (LRP) occurs in a trans-Golgi			
RT	compartment.";			
RL	EMBO J. 9:1769-1776(1990).			
RP	FUNCTION.			
EX	MEDLINE=91092405; PubMed=1702392;			
RA	Kristensen T., Moestrup S.K., Gliemann J., Bendtsen L., Sand O.,			
RA	Sottrup-Jensen L.;			
RT	"Evidence that the newly cloned low-density-lipoprotein receptor			

Q9Y561 homo sapien  
Q8IU80 homo sapien  
Q9DBI0 mus musculu  
P05156 homo sapien  
P98073 homo sapien  
P98153 homo sapien  
Q9YUW3 rattus norv  
P98074 sus scrofa  
P34434 caenorhabdi  
P14314 homo sapien  
Q61129 mus musculu  
P98072 bos taurus

[illegible]

```

Query Match          100.0%; Score 506; DB 1; Length 4544;
Best Local Similarity 100.0%; Pred. No. 6.2e-41;
Matches      86; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

QY      1 KTCSPKQFACRDITICISGWRCGRDGDGDEAPETCPQSKAQCQPNEHNCIGTEL 60
DB      25 KTCSPKQFACRDITICISGWRCGRDGDGDEAPETCPQSKAQCQPNEHNCIGTEL 84
      |||
QY      61 CVPMSRLCNGVQDCMDGSDGEGPHCRE 86
DB      85 CVPMSRLCNGVQDCMDGSDGEGPHCRE 110
      |||

RESULT 2
LRP1_CHICK STANDARD; PRT; 4543 AA.
ID      LRP1_CHICK
P98157;
01-OCT-1996 (Rel. 34, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
15-MAR-2004 (Rel. 43, Last annotation update)
Low-density lipoprotein receptor-related protein 1 precursor (LRP)
(Alpha-2-macroglobulin receptor) (A2MR).
Gallus gallus (Chicken).
OS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauia; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
Gallus.
NCBI_L TaxID=9031;
[1]
SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
STRAIN=White leghorn; TISSUE=Liver, and Ovary;
MEDLINE=941103212; PubMed=7506255;
Nimf J., Stifani S., Bailou P.T., Schneider W.J.;
"The somatic cell-specific low density lipoprotein receptor-related
protein of the chicken. Close kinship to mammalian low density
lipoprotein receptor gene family members.";
J. Biol. Chem. 269:212-219(1994).
-!- FUNCTION: Involved in the plasma clearance of chylomicron remnants
and activated alpha 2-macroglobulin, as well as the local
metabolism of complexes between plasminogen activators and their
endogenous inhibitors. Binds vitellogenin, calcium and alpha 2-
macroglobulin.
-!- SUBCELLULAR LOCATION: Type I membrane protein.
-!- ALTERNATIVE PRODUCTS.
Event=Alternative splicing; Named isoforms=2;
Name=1;
IsoId=P98157-1; Sequence=Displayed;
Name=2;
IsoId=P98157-2; Sequence=VSP_004312;
-!- TISSUE SPECIFICITY: Somatic.
-!- PTM: Cleaved into a 85 kDa membrane-spanning subunit (LRP-85) and
a 515 kDa large extracellular domain (LRP-515) that remains non-
covalently associated.
-!- SIMILARITY: Contains 22 EGF-like domains.
-!- SIMILARITY: Contains 31 LDL-receptor class A domains.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
-----
EMBL; X74904; CAA52870.1; -.
PIR; A53102; A53102.
PDB; 1LFX; 29-DEC-99.
InterPro; IPR000152; Asx_hydroxyl_S.
InterPro; IPR001861; EGF_Ca.
InterPro; IPR006209; EGF_like.
InterPro; IPR002172; LDL_receptor_A.
InterPro; IPR000033; LDL_receptor_rep.
Pfam; PF00008; EGF; 14.

```

FT	DOMAIN	4339	4374	EGF-LIKE 21.
FT	DOMAIN	4372	4409	EGF-LIKE 22.
FT	SITE	3939	3942	RECOGNITION SITE FOR PROTEOLYTICAL
FT				PROCESSING (POTENTIAL).
FT				ENDOCYTOSIS SIGNAL (POTENTIAL).
FT	SITE	4501	4506	BY SIMILARITY.
FT	DISULFID	29	42	BY SIMILARITY.
FT	DISULFID	36	55	BY SIMILARITY.
FT	DISULFID	49	66	BY SIMILARITY.
FT	DISULFID	74	87	BY SIMILARITY.
FT	DISULFID	81	100	BY SIMILARITY.
FT	DISULFID	94	110	BY SIMILARITY.
FT	DISULFID	117	126	BY SIMILARITY.
FT	DISULFID	122	135	BY SIMILARITY.
FT	DISULFID	137	150	BY SIMILARITY.
FT	DISULFID	156	166	BY SIMILARITY.
FT	DISULFID	162	175	BY SIMILARITY.
FT	DISULFID	177	190	BY SIMILARITY.
FT	DISULFID	480	495	BY SIMILARITY.
FT	DISULFID	491	506	BY SIMILARITY.
FT	DISULFID	508	521	BY SIMILARITY.
FT	DISULFID	805	816	BY SIMILARITY.
FT	DISULFID	812	825	BY SIMILARITY.
FT	DISULFID	827	840	BY SIMILARITY.
FT	DISULFID	852	864	BY SIMILARITY.
FT	DISULFID	859	877	BY SIMILARITY.
FT	DISULFID	871	888	BY SIMILARITY.
FT	DISULFID	893	905	BY SIMILARITY.
FT	DISULFID	900	918	BY SIMILARITY.
FT	DISULFID	912	929	BY SIMILARITY.
FT	DISULFID	934	946	BY SIMILARITY.
FT	DISULFID	941	959	BY SIMILARITY.
FT	DISULFID	953	969	BY SIMILARITY.
FT	DISULFID	974	987	BY SIMILARITY.
FT	DISULFID	982	1000	BY SIMILARITY.
FT	DISULFID	994	1009	BY SIMILARITY.
FT	DISULFID	1013	1025	BY SIMILARITY.
FT	DISULFID	1020	1038	BY SIMILARITY.
FT	DISULFID	1032	1049	BY SIMILARITY.
FT	DISULFID	1060	1073	BY SIMILARITY.
FT	DISULFID	1067	1086	BY SIMILARITY.
FT	DISULFID	1080	1095	BY SIMILARITY.
FT	DISULFID	1102	1116	BY SIMILARITY.
FT	DISULFID	1110	1129	BY SIMILARITY.
FT	DISULFID	1123	1138	BY SIMILARITY.
FT	DISULFID	1143	1157	BY SIMILARITY.
FT	DISULFID	1150	1170	BY SIMILARITY.
FT	DISULFID	1164	1180	BY SIMILARITY.
FT	DISULFID	1183	1194	BY SIMILARITY.
FT	DISULFID	1190	1204	BY SIMILARITY.
FT	DISULFID	1206	1219	BY SIMILARITY.
FT	DISULFID	1225	1235	BY SIMILARITY.
FT	DISULFID	1231	1244	BY SIMILARITY.
FT	DISULFID	1246	1259	BY SIMILARITY.
FT	DISULFID	1258	1271	BY SIMILARITY.
FT	DISULFID	1283	1296	BY SIMILARITY.
FT	DISULFID	1296	1309	BY SIMILARITY.
FT	DISULFID	1309	1322	BY SIMILARITY.
FT	DISULFID	1322	1335	BY SIMILARITY.
FT	DISULFID	1335	1348	BY SIMILARITY.
FT	DISULFID	1348	1361	BY SIMILARITY.
FT	DISULFID	1361	1374	BY SIMILARITY.
FT	DISULFID	1374	1387	BY SIMILARITY.
FT	DISULFID	1387	1400	BY SIMILARITY.
FT	DISULFID	1400	1413	BY SIMILARITY.
FT	DISULFID	1413	1426	BY SIMILARITY.
FT	DISULFID	1426	1439	BY SIMILARITY.
FT	DISULFID	1439	1452	BY SIMILARITY.
FT	DISULFID	1452	1465	BY SIMILARITY.
FT	DISULFID	1465	1478	BY SIMILARITY.
FT	DISULFID	1478	1491	BY SIMILARITY.
FT	DISULFID	1491	1504	BY SIMILARITY.
FT	DISULFID	1504	1517	BY SIMILARITY.
FT	DISULFID	1517	1530	BY SIMILARITY.
FT	DISULFID	1530	1543	BY SIMILARITY.
FT	DISULFID	1543	1556	BY SIMILARITY.
FT	DISULFID	1556	1569	BY SIMILARITY.
FT	DISULFID	1569	1582	BY SIMILARITY.
FT	DISULFID	1582	1595	BY SIMILARITY.
FT	DISULFID	1595	1608	BY SIMILARITY.
FT	DISULFID	1608	1621	BY SIMILARITY.
FT	DISULFID	1621	1634	BY SIMILARITY.
FT	DISULFID	1634	1647	BY SIMILARITY.
FT	DISULFID	1647	1660	BY SIMILARITY.
FT	DISULFID	1660	1673	BY SIMILARITY.
FT	DISULFID	1673	1686	BY SIMILARITY.
FT	DISULFID	1686	1699	BY SIMILARITY.
FT	DISULFID	1699	1712	BY SIMILARITY.
FT	DISULFID	1712	1725	BY SIMILARITY.
FT	DISULFID	1725	1738	BY SIMILARITY.
FT	DISULFID	1738	1751	BY SIMILARITY.
FT	DISULFID	1751	1764	BY SIMILARITY.
FT	DISULFID	1764	1777	BY SIMILARITY.
FT	DISULFID	1777	1790	BY SIMILARITY.
FT	DISULFID	1790	1803	BY SIMILARITY.
FT	DISULFID	1803	1816	BY SIMILARITY.
FT	DISULFID	1816	1829	BY SIMILARITY.
FT	DISULFID	1829	1842	BY SIMILARITY.
FT	DISULFID	1842	1855	BY SIMILARITY.
FT	DISULFID	1855	1868	BY SIMILARITY.
FT	DISULFID	1868	1881	BY SIMILARITY.
FT	DISULFID	1881	1894	BY SIMILARITY.
FT	DISULFID	1894	1907	BY SIMILARITY.
FT	DISULFID	1907	1920	BY SIMILARITY.
FT	DISULFID	1920	1933	BY SIMILARITY.
FT	DISULFID	1933	1946	BY SIMILARITY.
FT	DISULFID	1946	1959	BY SIMILARITY.
FT	DISULFID	1959	1972	BY SIMILARITY.
FT	DISULFID	1972	1985	BY SIMILARITY.
FT	DISULFID	1985	1998	BY SIMILARITY.
FT	DISULFID	1998	2011	BY SIMILARITY.
FT	DISULFID	2011	2024	BY SIMILARITY.
FT	DISULFID	2024	2037	BY SIMILARITY.
FT	DISULFID	2037	2050	BY SIMILARITY.
FT	DISULFID	2050	2063	BY SIMILARITY.
FT	DISULFID	2063	2076	BY SIMILARITY.
FT	DISULFID	2076	2089	BY SIMILARITY.
FT	DISULFID	2089	2102	BY SIMILARITY.
FT	DISULFID	2102	2115	BY SIMILARITY.
FT	DISULFID	2115	2128	BY SIMILARITY.
FT	DISULFID	2128	2141	BY SIMILARITY.
FT	DISULFID	2141	2154	BY SIMILARITY.
FT	DISULFID	2154	2167	BY SIMILARITY.
FT	DISULFID	2167	2180	BY SIMILARITY.
FT	DISULFID	2180	2193	BY SIMILARITY.
FT	DISULFID	2193	2206	BY SIMILARITY.
FT	DISULFID	2206	2219	BY SIMILARITY.
FT	DISULFID	2219	2232	BY SIMILARITY.
FT	DISULFID	2232	2245	BY SIMILARITY.
FT	DISULFID	2245	2258	BY SIMILARITY.
FT	DISULFID	2258	2271	BY SIMILARITY.
FT	DISULFID	2271	2284	BY SIMILARITY.
FT	DISULFID	2284	2297	BY SIMILARITY.
FT	DISULFID	2297	2310	BY SIMILARITY.
FT	DISULFID	2310	2323	BY SIMILARITY.
FT	DISULFID	2323	2336	BY SIMILARITY.
FT	DISULFID	2336	2349	BY SIMILARITY.
FT	DISULFID	2349	2362	BY SIMILARITY.
FT	DISULFID	2362	2375	BY SIMILARITY.
FT	DISULFID	2375	2388	BY SIMILARITY.
FT	DISULFID	2388	2401	BY SIMILARITY.
FT	DISULFID	2401	2414	BY SIMILARITY.
FT	DISULFID	2414	2427	BY SIMILARITY.
FT	DISULFID	2427	2440	BY SIMILARITY.
FT	DISULFID	2440	2453	BY SIMILARITY.
FT	DISULFID	2453	2466	BY SIMILARITY.
FT	DISULFID	2466	2479	BY SIMILARITY.
FT	DISULFID	2479	2492	BY SIMILARITY.
FT	DISULFID	2492	2505	BY SIMILARITY.
FT	DISULFID	2505	2518	BY SIMILARITY.
FT	DISULFID	2518	2531	BY SIMILARITY.
FT	DISULFID	2531	2544	BY SIMILARITY.
FT	DISULFID	2544	2557	BY SIMILARITY.
FT	DISULFID	2557	2570	BY SIMILARITY.
FT	DISULFID	2570	2583	BY SIMILARITY.
FT	DISULFID	2583	2596	BY SIMILARITY.
FT	DISULFID	2596	2609	BY SIMILARITY.
FT	DISULFID	2609	2622	BY SIMILARITY.
FT	DISULFID	2622	2635	BY SIMILARITY.
FT	DISULFID	2635	2648	BY SIMILARITY.
FT	DISULFID	2648	2661	BY SIMILARITY.
FT	DISULFID	2661	2674	BY SIMILARITY.
FT	DISULFID	2674	2687	BY SIMILARITY.
FT	DISULFID	2687	2700	BY SIMILARITY.
FT	DISULFID	2700	2713	BY SIMILARITY.
FT	DISULFID	2713	2726	BY SIMILARITY.
FT	DISULFID	2726	2739	BY SIMILARITY.
FT	DISULFID	2739	2752	BY SIMILARITY.
FT	DISULFID	2752	2765	BY SIMILARITY.
FT	DISULFID	2765	2778	BY SIMILARITY.
FT	DISULFID	2778	2791	BY SIMILARITY.
FT	DISULFID	2791	2804	BY SIMILARITY.
FT	DISULFID	2804	2817	BY SIMILARITY.
FT	DISULFID	2817	2830	BY SIMILARITY.
FT	DISULFID	2830	2843	BY SIMILARITY.
FT	DISULFID	2843	2856	BY SIMILARITY.
FT	DISULFID	2856	2869	BY SIMILARITY.
FT	DISULFID	2869	2882	BY SIMILARITY.
FT	DISULFID	2882	2895	BY SIMILARITY.
FT	DISULFID	2895	2908	BY SIMILARITY.
FT	DISULFID	2908	2921	BY SIMILARITY.
FT	DISULFID	2921	2934	BY SIMILARITY.
FT	DISULFID	2934	2947	BY SIMILARITY.
FT	DISULFID	2947	2960	BY SIMILARITY.
FT	DISULFID	2960	2973	BY SIMILARITY.
FT	DISULFID	2973	2986	BY SIMILARITY.
FT	DISULFID	2986	2999	BY SIMILARITY.
FT	DISULFID	2999	3012	BY SIMILARITY.
FT	DISULFID	3012	3025	BY SIMILARITY.
FT	DISULFID	3025	3038	BY SIMILARITY.
FT	DISULFID	3038	3051	BY SIMILARITY.
FT	DISULFID	3051	3064	BY SIMILARITY.
FT	DISULFID	3064	3077	BY SIMILARITY.
FT	DISULFID	3077	3090	BY SIMILARITY.
FT	DISULFID	3090	3103	BY SIMILARITY.
FT	DISULFID	3103	3116	BY SIMILARITY.
FT	DISULFID	3116	3129	BY SIMILARITY.
FT	DISULFID	3129	3142	BY SIMILARITY.
FT	DISULFID	3142	3155	BY SIMILARITY.
FT	DISULFID	3155	3168	BY SIMILARITY.
FT	DISULFID	3168	3181	BY SIMILARITY.
FT	DISULFID	3181	3194	BY SIMILARITY.
FT	DISULFID	3194	3207	BY SIMILARITY.
FT	DISULFID	3207	3220	BY SIMILARITY.
FT	DISULFID	3220	3233	BY SIMILARITY.
FT	DISULFID	3233	3246	BY SIMILARITY.
FT	DISULFID	3246	3259	BY SIMILARITY.
FT	DISULFID	3259	3272	BY SIMILARITY.
FT	DISULFID	3272	3285	BY SIMILARITY.
FT	DISULFID	3285	3298	BY SIMILARITY.
FT	DISULFID	3298	3311	BY SIMILARITY.
FT	DISULFID	3311	3324	BY SIMILARITY.
FT	DISULFID	3324	3337	BY SIMILARITY.
FT	DISULFID	3337	3350	BY SIMILARITY.
FT	DISULFID	3350	3363	BY SIMILARITY.
FT	DISULFID	3363	3376	BY SIMILARITY.
FT	DISULFID	3376	3389	BY SIMILARITY.
FT	DISULFID	3389	3402	BY SIMILARITY.
FT	DISULFID	3402	3415	BY SIMILARITY.
FT	DISULFID	3415	3428	BY SIMILARITY.
FT	DISULFID	3428	3441	BY SIMILARITY.
FT	DISULFID	3441	3454	BY SIMILARITY.
FT	DISULFID	3454	3467	BY SIMILARITY.
FT	DISULFID	3467	3480	BY SIMILARITY.
FT	DISULFID	3480	3493	BY SIMILARITY.
FT	DISULFID	3493	3506	BY SIMILARITY.
FT	DISULFID	3506	3519	BY SIMILARITY.
FT	DISULFID	3519	3532	BY SIMILARITY.
FT	DISULFID	3532	3545	BY SIMILARITY.
FT	DISULFID	3545	3558	BY SIMILARITY.
FT	DISULFID	3558	3571	BY SIMILARITY.
FT	DISULFID	3571	3584	BY SIMILARITY.
FT	DISULFID	3584	3597	BY SIMILARITY.
FT	DISULFID	3597	3610	BY SIMILARITY.
FT	DISULFID	3610	3623	BY SIMILARITY.
FT	DISULFID	3623	3636	BY SIMILARITY.
FT	DISULFID	3636	3649	BY SIMILARITY.
FT	DISULFID	3649	3662	BY SIMILARITY.
FT	DISULFID	3662	3675	BY SIMILARITY.
FT	DISULFID	3675	3688	BY SIMILARITY.
FT	DISULFID	3688	3701	BY SIMILARITY.
FT	DISULFID	3701	3714	BY SIMILARITY.
FT	DISULFID	3714	3727	BY SIMILARITY.
FT	DISULFID	3727	3740	BY SIMILARITY.
FT	DISULFID	3740	3753	BY SIMILARITY.
FT	DISULFID	3753	3766	BY SIMILARITY.
FT	DISULFID	3766	3779	BY SIMILARITY.
FT	DISULFID	3779	3792	BY SIMILARITY.
FT	DISULFID	3792	3805	BY SIMILARITY.
FT	DISULFID	3805	3818	BY SIMILARITY.
FT	DISULFID	3818	3831	BY SIMILARITY.
FT	DISULFID	3831	3844	BY SIMILARITY.
FT	DISULFID	3844	3857	BY SIMILARITY.
FT	DISULFID	3857	3870	BY SIMILARITY.
FT	DISULFID	3870	3883	BY SIMILARITY.
FT	DISULFID	3883	3896	BY SIMILARITY.
FT	DISULFID	3896	3909	BY SIMILARITY.
FT	DISULFID	3909	3922	BY SIMILARITY.
FT	DISULFID	3922	3935	BY SIMILARITY.
FT	DISULFID	3935	3948	BY SIMILARITY.
FT	DISULFID	3948	3961	BY SIMILARITY.
FT	DISULFID	3961	3974	BY SIMIL

[illegible]

Query Match 40.4%; Score 204.5; DB 1; Length 4753;  
Best Local Similarity 36.0%; Pred. No. 7.1e-12;  
Matches 41; Conservative 10; Mismatches 26; Indels 37;

3 CSPKQFACRDQITCISKGWRCDGERDCPDGSDEAPEICP-----QSKAQRCQPNEH 53 QY

```
Db 182 CFQYQFCADKTKQIQKSWVCDGSKDCADGSD-EDTCEFFKKTANEFQCKNKRQPKRF 240
QY 54 -----NCLGTCLVPMNSRLNGVQDCWDGSDSE 80
Db 241 RCDYDDCGNSDEDECGYRCPGKWNCPGTHCHDIDQLKLCDSKOCADGADE 294

RESULT 4
LRP2_RAT
ID _LRP2_RAT STANDARD; PRT; 4660 AA.
AC P98158;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Low-density lipoprotein receptor-related protein 2 precursor (Megalyn)
DE (Glycoprotein 330) (gp330).
GN LRP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Kidney;
RX MEDLINE=95024033; PubMed=7937880;
RA Saito A., Pietromonaco S., Loo A.K.C., Farquhar M.G.;
RT "Complete cloning and sequencing of rat gp330/'megalyn,' a
RT distinctive member of the low density lipoprotein receptor gene
RT family.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:9725-9729(1994).
RN [2]
RP FUNCTION.
RX MEDLINE=95386696; PubMed=7544804;
RA Moestrup S.K., Cui S., Vorum H., Bregengaard C., Bjorn S.E.,
RA Norris K., Gliemann J., Christensen E.I.;
RT "Evidence that epithelial glycoprotein 330/megalyn mediates uptake of
RT polybasic drugs.";
RL J. Clin. Invest. 96:1404-1413(1995).
RN [3]
RP TISSUE SPECIFICITY.
RX MEDLINE=94172242; PubMed=7510321;
RA Zheng G., Bachinsky D.R., Stamenkovic I., Strickland D.K., Brown D.,
RA Andres G., McCluskey R.T.;
RT "Organ distribution in rats of two members of the low-density
RT lipoprotein receptor gene family, gp330 and LRP/alpha 2MR, and the
RT receptor-associated protein (RAP).";
RL J. Histochem. Cytochem. 42:531-542(1994).
CC -I- FUNCTION: Binds plasminogen, extracellular matrix components,
CC plasminogen activator-plasminogen activator inhibitor type I
CC complex, apolipoprotein E-enriched beta-VLDL, lipoprotein lipase,
CC lactoferrin, clusterin and calcium.
CC -I- FUNCTION: Receptor-mediated uptake of polybasic drugs such as
CC aprotinin, aminoglycosides and polymyxin B.
CC -I- SUBUNIT: Forms a multimeric complex together with a receptor-
CC associated protein (RAP). Binds to ankyrin-repeat family A protein
CC 2 (ANKRA2) (By similarity).
CC -I- SUBCELLULAR LOCATION: Type I membrane protein. Expressed in
CC clathrin-coated pits; a soluble form is possibly derived by
CC cleavage at the cell surface.
CC -I- TISSUE SPECIFICITY: Epithelial cells of kidney glomerulus and
CC proximal tubule, lung, epididymis, yolk sac, among others.
CC -I- SIMILARITY: Contains 36 LDL-receptor class A domains.
CC -I- SIMILARITY: Contains 37 LDL-receptor class B domains.
CC -I- SIMILARITY: Contains 17 EGF-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sb.ch)
```

```
CC EMBL; L34049; AAA51369.1; -.
DR PIR; T42737; T42737.
DR HSSP; Q07954; ICR8.
DR GlycosultedB; P98158; -.
DR InterPro; IPR000152; Asx_hydroxyl_S.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000033; Ldl_receptor_rep.
DR Pfam; PF00008; EGF; 9.
DR Pfam; PF00057; ldl_recept_a; 36.
DR Pfam; PF00058; ldl_recept_b; 33.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00179; EGF_CA; 3.
DR SMART; SM00192; LDLA; 36.
DR SMART; SM00135; LY; 34.
DR PROSITE; PS00010; ASX_HYDROXYL; 4.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 8.
DR PROSITE; PS00026; EGF_3; 8.
DR PROSITE; PS01187; EGF_CA; 3.
DR PROSITE; PS01209; LDLRA_1; 31.
DR PROSITE; PS00068; LDLRA_2; 36.
KW Glycoprotein; Repeat; Endocytosis; Coated pits; Transmembrane;
KW Receptor; EGF-like domain; SH3-binding; Signal.
FT SIGNAL 1 25
FT CHAIN 26 4660
FT DOMAIN 26 4425
FT TRANSMEM 4426 4446
FT DOMAIN 4447 4660
FT DOMAIN 26 64
FT DOMAIN 65 105
FT DOMAIN 106 144
FT DOMAIN 145 181
FT DOMAIN 181 219
FT DOMAIN 220 258
FT DOMAIN 263 307
FT DOMAIN 308 346
FT DOMAIN 347 385
FT DOMAIN 435 476
FT DOMAIN 478 519
FT DOMAIN 521 566
FT DOMAIN 568 611
FT DOMAIN 612 652
FT DOMAIN 658 704
FT DOMAIN 752 793
FT DOMAIN 837 879
FT DOMAIN 881 923
FT DOMAIN 969 1013
FT DOMAIN 1023 1061
FT DOMAIN 1064 1103
FT DOMAIN 1108 1146
FT DOMAIN 1148 1186
FT DOMAIN 1186 1225
FT DOMAIN 1229 1269
FT DOMAIN 1270 1308
FT DOMAIN 1311 1351
FT DOMAIN 1350 1390
FT DOMAIN 1391 1430
FT DOMAIN 1479 1520
FT DOMAIN 1522 1563
FT DOMAIN 1567 1609
FT DOMAIN 1611 1654
FT DOMAIN 1656 1696
FT DOMAIN 1701 1742
FT DOMAIN 1791 1832
FT DOMAIN 1834 1882
FT DOMAIN 1884 1930
FT DOMAIN 1932 1972
FT DOMAIN 1973 2013
FT DOMAIN 2013 2053
FT DOMAIN 2053 2093
FT DOMAIN 2093 2133
FT DOMAIN 2133 2173
FT DOMAIN 2173 2213
FT DOMAIN 2213 2253
FT DOMAIN 2253 2293
FT DOMAIN 2293 2333
FT DOMAIN 2333 2373
FT DOMAIN 2373 2413
FT DOMAIN 2413 2453
FT DOMAIN 2453 2493
FT DOMAIN 2493 2533
FT DOMAIN 2533 2573
FT DOMAIN 2573 2613
FT DOMAIN 2613 2653
FT DOMAIN 2653 2693
FT DOMAIN 2693 2733
FT DOMAIN 2733 2773
FT DOMAIN 2773 2813
FT DOMAIN 2813 2853
FT DOMAIN 2853 2893
FT DOMAIN 2893 2933
FT DOMAIN 2933 2973
FT DOMAIN 2973 3013
FT DOMAIN 3013 3053
FT DOMAIN 3053 3093
FT DOMAIN 3093 3133
FT DOMAIN 3133 3173
FT DOMAIN 3173 3213
FT DOMAIN 3213 3253
FT DOMAIN 3253 3293
FT DOMAIN 3293 3333
FT DOMAIN 3333 3373
FT DOMAIN 3373 3413
FT DOMAIN 3413 3453
FT DOMAIN 3453 3493
FT DOMAIN 3493 3533
FT DOMAIN 3533 3573
FT DOMAIN 3573 3613
FT DOMAIN 3613 3653
FT DOMAIN 3653 3693
FT DOMAIN 3693 3733
FT DOMAIN 3733 3773
FT DOMAIN 3773 3813
FT DOMAIN 3813 3853
FT DOMAIN 3853 3893
FT DOMAIN 3893 3933
FT DOMAIN 3933 3973
FT DOMAIN 3973 4013
FT DOMAIN 4013 4053
FT DOMAIN 4053 4093
FT DOMAIN 4093 4133
FT DOMAIN 4133 4173
FT DOMAIN 4173 4213
FT DOMAIN 4213 4253
FT DOMAIN 4253 4293
FT DOMAIN 4293 4333
FT DOMAIN 4333 4373
FT DOMAIN 4373 4413
FT DOMAIN 4413 4453
FT DOMAIN 4453 4493
FT DOMAIN 4493 4533
FT DOMAIN 4533 4573
FT DOMAIN 4573 4613
FT DOMAIN 4613 4653
FT DOMAIN 4653 4693
FT DOMAIN 4693 4733
FT DOMAIN 4733 4773
FT DOMAIN 4773 4813
FT DOMAIN 4813 4853
FT DOMAIN 4853 4893
FT DOMAIN 4893 4933
FT DOMAIN 4933 4973
FT DOMAIN 4973 5013
FT DOMAIN 5013 5053
FT DOMAIN 5053 5093
FT DOMAIN 5093 5133
FT DOMAIN 5133 5173
FT DOMAIN 5173 5213
FT DOMAIN 5213 5253
FT DOMAIN 5253 5293
FT DOMAIN 5293 5333
FT DOMAIN 5333 5373
FT DOMAIN 5373 5413
FT DOMAIN 5413 5453
FT DOMAIN 5453 5493
FT DOMAIN 5493 5533
FT DOMAIN 5533 5573
FT DOMAIN 5573 5613
FT DOMAIN 5613 5653
FT DOMAIN 5653 5693
FT DOMAIN 5693 5733
FT DOMAIN 5733 5773
FT DOMAIN 5773 5813
FT DOMAIN 5813 5853
FT DOMAIN 5853 5893
FT DOMAIN 5893 5933
FT DOMAIN 5933 5973
FT DOMAIN 5973 6013
FT DOMAIN 6013 6053
FT DOMAIN 6053 6093
FT DOMAIN 6093 6133
FT DOMAIN 6133 6173
FT DOMAIN 6173 6213
FT DOMAIN 6213 6253
FT DOMAIN 6253 6293
FT DOMAIN 6293 6333
FT DOMAIN 6333 6373
FT DOMAIN 6373 6413
FT DOMAIN 6413 6453
FT DOMAIN 6453 6493
FT DOMAIN 6493 6533
FT DOMAIN 6533 6573
FT DOMAIN 6573 6613
FT DOMAIN 6613 6653
FT DOMAIN 6653 6693
FT DOMAIN 6693 6733
FT DOMAIN 6733 6773
FT DOMAIN 6773 6813
FT DOMAIN 6813 6853
FT DOMAIN 6853 6893
FT DOMAIN 6893 6933
FT DOMAIN 6933 6973
FT DOMAIN 6973 7013
FT DOMAIN 7013 7053
FT DOMAIN 7053 7093
FT DOMAIN 7093 7133
FT DOMAIN 7133 7173
FT DOMAIN 7173 7213
FT DOMAIN 7213 7253
FT DOMAIN 7253 7293
FT DOMAIN 7293 7333
FT DOMAIN 7333 7373
FT DOMAIN 7373 7413
FT DOMAIN 7413 7453
FT DOMAIN 7453 7493
FT DOMAIN 7493 7533
FT DOMAIN 7533 7573
FT DOMAIN 7573 7613
FT DOMAIN 7613 7653
FT DOMAIN 7653 7693
FT DOMAIN 7693 7733
FT DOMAIN 7733 7773
FT DOMAIN 7773 7813
FT DOMAIN 7813 7853
FT DOMAIN 7853 7893
FT DOMAIN 7893 7933
FT DOMAIN 7933 7973
FT DOMAIN 7973 8013
FT DOMAIN 8013 8053
FT DOMAIN 8053 8093
FT DOMAIN 8093 8133
FT DOMAIN 8133 8173
FT DOMAIN 8173 8213
FT DOMAIN 8213 8253
FT DOMAIN 8253 8293
FT DOMAIN 8293 8333
FT DOMAIN 8333 8373
FT DOMAIN 8373 8413
FT DOMAIN 8413 8453
FT DOMAIN 8453 8493
FT DOMAIN 8493 8533
FT DOMAIN 8533 8573
FT DOMAIN 8573 8613
FT DOMAIN 8613 8653
FT DOMAIN 8653 8693
FT DOMAIN 8693 8733
FT DOMAIN 8733 8773
FT DOMAIN 8773 8813
FT DOMAIN 8813 8853
FT DOMAIN 8853 8893
FT DOMAIN 8893 8933
FT DOMAIN 8933 8973
FT DOMAIN 8973 9013
FT DOMAIN 9013 9053
FT DOMAIN 9053 9093
FT DOMAIN 9093 9133
FT DOMAIN 9133 9173
FT DOMAIN 9173 9213
FT DOMAIN 9213 9253
FT DOMAIN 9253 9293
FT DOMAIN 9293 9333
FT DOMAIN 9333 9373
FT DOMAIN 9373 9413
FT DOMAIN 9413 9453
FT DOMAIN 9453 9493
FT DOMAIN 9493 9533
FT DOMAIN 9533 9573
FT DOMAIN 9573 9613
FT DOMAIN 9613 9653
FT DOMAIN 9653 9693
FT DOMAIN 9693 9733
FT DOMAIN 9733 9773
FT DOMAIN 9773 9813
FT DOMAIN 9813 9853
FT DOMAIN 9853 9893
FT DOMAIN 9893 9933
FT DOMAIN 9933 9973
FT DOMAIN 9973 10013
FT DOMAIN 10013 10053
FT DOMAIN 10053 10093
FT DOMAIN 10093 10133
FT DOMAIN 10133 10173
FT DOMAIN 10173 10213
FT DOMAIN 10213 10253
FT DOMAIN 10253 10293
FT DOMAIN 10293 10333
FT DOMAIN 10333 10373
FT DOMAIN 10373 10413
FT DOMAIN 10413 10453
FT DOMAIN 10453 10493
FT DOMAIN 10493 10533
FT DOMAIN 10533 10573
FT DOMAIN 10573 10613
FT DOMAIN 10613 10653
FT DOMAIN 10653 10693
FT DOMAIN 10693 10733
FT DOMAIN 10733 10773
FT DOMAIN 10773 10813
FT DOMAIN 10813 10853
FT DOMAIN 10853 10893
FT DOMAIN 10893 10933
FT DOMAIN 10933 10973
FT DOMAIN 10973 11013
FT DOMAIN 11013 11053
FT DOMAIN 11053 11093
FT DOMAIN 11093 11133
FT DOMAIN 11133 11173
FT DOMAIN 11173 11213
FT DOMAIN 11213 11253
FT DOMAIN 11253 11293
FT DOMAIN 11293 11333
FT DOMAIN 11333 11373
FT DOMAIN 11373 11413
FT DOMAIN 11413 11453
FT DOMAIN 11453 11493
FT DOMAIN 11493 11533
FT DOMAIN 11533 11573
FT DOMAIN 11573 11613
FT DOMAIN 11613 11653
FT DOMAIN 11653 11693
FT DOMAIN 11693 11733
FT DOMAIN 11733 11773
FT DOMAIN 11773 11813
FT DOMAIN 11813 11853
FT DOMAIN 11853 11893
FT DOMAIN 11893 11933
FT DOMAIN 11933 11973
FT DOMAIN 11973 12013
FT DOMAIN 12013 12053
FT DOMAIN 12053 12093
FT DOMAIN 12093 12133
FT DOMAIN 12133 12173
FT DOMAIN 12173 12213
FT DOMAIN 12213 12253
FT DOMAIN 12253 12293
FT DOMAIN 12293 12333
FT DOMAIN 12333 12373
FT DOMAIN 12373 12413
FT DOMAIN 12413 12453
FT DOMAIN 12453 12493
FT DOMAIN 12493 12533
FT DOMAIN 12533 12573
FT DOMAIN 12573 12613
FT DOMAIN 12613 12653
FT DOMAIN 12653 12693
FT DOMAIN 12693 12733
FT DOMAIN 12733 12773
FT DOMAIN 12773 12813
FT DOMAIN 12813 12853
FT DOMAIN 12853 12893
FT DOMAIN 12893 12933
FT DOMAIN 12933 12973
FT DOMAIN 12973 13013
FT DOMAIN 13013 13053
FT DOMAIN 13053 13093
FT DOMAIN 13093 13133
FT DOMAIN 13133 13173
FT DOMAIN 13173 13213
FT DOMAIN 13213 13253
FT DOMAIN 13253 13293
FT DOMAIN 13293 13333
FT DOMAIN 13333 13373
FT DOMAIN 13373 13413
FT DOMAIN 13413 13453
FT DOMAIN 13453 13493
FT DOMAIN 13493 13533
FT DOMAIN 13533 13573
FT DOMAIN 13573 13613
FT DOMAIN 13613 13653
FT DOMAIN 13653 13693
FT DOMAIN 13693 13733
FT DOMAIN 13733 13773
FT DOMAIN 13773 13813
FT DOMAIN 13813 13853
FT DOMAIN 13853 13893
FT DOMAIN 13893 13933
FT DOMAIN 13933 13973
FT DOMAIN 13973 14013
FT DOMAIN 14013 14053
FT DOMAIN 14053 14093
FT DOMAIN 14093 14133
FT DOMAIN 14133 14173
FT DOMAIN 14173 14213
FT DOMAIN 14213 14253
FT DOMAIN 14253 14293
FT DOMAIN 14293 14333
FT DOMAIN 14333 14373
FT DOMAIN 14373 14413
FT DOMAIN 14413 14453
FT DOMAIN 14453 14493
FT DOMAIN 14493 14533
FT DOMAIN 14533 14573
FT DOMAIN 14573 14613
FT DOMAIN 14613 14653
FT DOMAIN 14653 14693
FT DOMAIN 14693 14733
FT DOMAIN 14733 14773
FT DOMAIN 14773 14813
FT DOMAIN 14813 14853
FT DOMAIN 14853 14893
FT DOMAIN 14893 14933
FT DOMAIN 14933 14973
FT DOMAIN 14973 15013
FT DOMAIN 15013 15053
FT DOMAIN 15053 15093
FT DOMAIN 15093 15133
FT DOMAIN 15133 15173
FT DOMAIN 15173 15213
FT DOMAIN 15213 15253
FT DOMAIN 15253 15293
FT DOMAIN 15293 15333
FT DOMAIN 15333 15373
FT DOMAIN 15373 15413
FT DOMAIN 15413 15453
FT DOMAIN 15453 15493
FT DOMAIN 15493 15533
FT DOMAIN 15533 15573
FT DOMAIN 15573 15613
FT DOMAIN 15613 15653
FT DOMAIN 15653 15693
FT DOMAIN 15693 15733
FT DOMAIN 15733 15773
FT DOMAIN 15773 15813
FT DOMAIN 15813 15853
FT DOMAIN 15853 15893
FT DOMAIN 15893 15933
FT DOMAIN 15933 15973
FT DOMAIN 15973 16013
FT DOMAIN 16013 16053
FT DOMAIN 16053 16093
FT DOMAIN 16093 16133
FT DOMAIN 16133 16173
FT DOMAIN 16173 16213
FT DOMAIN 16213 16253
FT DOMAIN 16253 16293
FT DOMAIN 16293 16333
FT DOMAIN 16333 16373
FT DOMAIN 16373 16413
FT DOMAIN 16413 16453
FT DOMAIN 16453 16493
FT DOMAIN 16493 16533
FT DOMAIN 16533 16573
FT DOMAIN 16573 16613
FT DOMAIN 16613 16653
FT DOMAIN 16653 16693
FT DOMAIN 16693 16733
FT DOMAIN 16733 16773
FT DOMAIN 16773 16813
FT DOMAIN 16813 16853
FT DOMAIN 16853 16893
FT DOMAIN 16893 16933
FT DOMAIN 16933 16973
FT DOMAIN 16973 17013
FT DOMAIN 17013 17053
FT DOMAIN 17053 17093
FT DOMAIN 17093 17133
FT DOMAIN 17133 17173
FT DOMAIN 17173 17213
FT DOMAIN 17213 17253
FT DOMAIN 17253 17293
FT DOMAIN 17293 17333
FT DOMAIN 17333 17373
FT DOMAIN 17373 17413
FT DOMAIN 17413 17453
FT DOMAIN 17453 17493
FT DOMAIN 17493 17533
FT DOMAIN 17533 17573
FT DOMAIN 17573 17613
FT DOMAIN 17613 17653
FT DOMAIN 17653 17693
FT DOMAIN 17693 17733
FT DOMAIN 17733 17773
FT DOMAIN 17773 17813
FT DOMAIN 17813 17853
FT DOMAIN 17853 17893
FT DOMAIN 17893 17933
FT DOMAIN 17933 17973
FT DOMAIN 17973 18013
FT DOMAIN 18013 18053
FT DOMAIN 18053 18093
FT DOMAIN 18093 18133
FT DOMAIN 18133 18173
FT DOMAIN 18173 18213
FT DOMAIN 18213 18253
FT DOMAIN 18253 18293
FT DOMAIN 18293 18333
FT DOMAIN 18333 18373
FT DOMAIN 18373 18413
FT DOMAIN 18413 18453
FT DOMAIN 18453 18493
FT DOMAIN 18493 18533
FT DOMAIN 18533 18573
FT DOMAIN 18573 18613
FT DOMAIN 18613 18653
FT DOMAIN 18653 18693
FT DOMAIN 18693 18733
FT DOMAIN 18733 18773
FT DOMAIN 18773 18813
FT DOMAIN 18813 18853
FT DOMAIN 18853 18893
FT DOMAIN 18893 18933
FT DOMAIN 18933 18973
FT DOMAIN 18973 19013
FT DOMAIN 19013 19053
FT DOMAIN 19053 19093
FT DOMAIN 19093 19133
FT DOMAIN 19133 19173
FT DOMAIN 19173 19213
FT DOMAIN 19213 19253
FT DOMAIN 19253 19293
FT DOMAIN 19293 19333
FT DOMAIN 19333 19373
FT DOMAIN 19373 19413
FT DOMAIN 19413 19453
FT DOMAIN 19453 19493
FT DOMAIN 19493 19533
FT DOMAIN 19533 19573
FT DOMAIN 19573 19613
FT DOMAIN 19613 19653
FT DOMAIN 19653 19693
FT DOMAIN 19693 19733
FT DOMAIN 19733 19773
FT DOMAIN 19773 19813
FT DOMAIN 19813 19853
FT DOMAIN 19853 19893
FT DOMAIN 19893 19933
FT DOMAIN 19933 19973
FT DOMAIN 19973 20013
FT DOMAIN 20013 20053
FT DOMAIN 20053 20093
FT DOMAIN 20093 20133
FT DOMAIN 20133 20173
FT DOMAIN 20173 20213
FT DOMAIN 20213 20253
FT DOMAIN 20253 20293
FT DOMAIN 20293 20333
FT DOMAIN 20333 20373
FT DOMAIN 20373 20413
FT DOMAIN 20413 20453
FT DOMAIN 20453 20493
FT DOMAIN 20493 20533
FT DOMAIN 20533 20573
FT DOMAIN 20573 20613
FT DOMAIN 20613 20653
FT DOMAIN 20653 20693
FT DOMAIN 20693 20733
FT DOMAIN 20733 20773
FT DOMAIN 20773 20813
FT DOMAIN 20813 20853
FT DOMAIN 20853 20893
FT DOMAIN 20893 20933
FT DOMAIN 20933 20973
FT DOMAIN 20973 21013
FT DOMAIN 21013 21053
FT DOMAIN 21053 21093
FT DOMAIN 21093 21133
FT DOMAIN 21133 21173
FT DOMAIN 21173 21213
FT DOMAIN 21213 21253
FT DOMAIN 21253 21293
FT DOMAIN 21293 21333
FT DOMAIN 21333 21373
FT DOMAIN 21373 21413
FT DOMAIN 21413 21453
FT DOMAIN 21453 21493
FT DOMAIN 21493 21533
FT DOMAIN 21533 21573
FT DOMAIN 21573 21613
FT DOMAIN 21613 21653
FT DOMAIN 21653 21693
FT DOMAIN 21693 21733
FT DOMAIN 21733 21773
FT DOMAIN 21773 21813
FT DOMAIN 21813 21853
FT DOMAIN 21853 21893
FT DOMAIN 21893 21933
FT DOMAIN 21933 21973
FT DOMAIN 21973 22013
FT DOMAIN 22013 22053
FT DOMAIN 22053 22093
FT DOMAIN 22093 22133
FT DOMAIN 22133 22173
FT DOMAIN 22173 22213
FT DOMAIN 22213 22253
FT DOMAIN 22253 22293
FT DOMAIN 22293 22333
FT DOMAIN 22333 22373
FT DOMAIN 22373 22413
FT DOMAIN 22413 22453
FT DOMAIN 22453 22493
FT DOMAIN 22493 22533
FT DOMAIN 22533 22573
FT DOMAIN 22573 22613
FT DOMAIN 22613 22653
FT DOMAIN 22653 22693
FT DOMAIN 22693 22733
FT DOMAIN 22733 22773
FT DOMAIN 22773 22813
FT DOMAIN 22813 22853
FT DOMAIN 22853 22893
FT DOMAIN 22893 22933
FT DOMAIN 22933 22973
FT DOMAIN 22973 23013
FT DOMAIN 23013 23053
FT DOMAIN 23053 23093
FT DOMAIN 23093 23133
FT DOMAIN 23133 23173
FT DOMAIN 23173 23213
FT DOMAIN 23213 23253
FT DOMAIN 23253 23293
FT DOMAIN 23293 23333
FT DOMAIN 23333 23373
FT DOMAIN 23373 23413
FT DOMAIN 23413 23453
FT DOMAIN 23453 23493
FT DOMAIN 23493 23533
FT DOMAIN 23533 23573
FT DOMAIN 23573 23613
FT DOMAIN 23613 23653
FT DOMAIN 23653 23693
FT DOMAIN 23693 23733
FT DOMAIN 23733 23773
FT DOMAIN 23773 23813
FT DOMAIN 23813 23853
FT DOMAIN 23853 23893
FT DOMAIN 23893 23933
FT DOMAIN 23933 23973
FT DOMAIN 23973 24013
FT DOMAIN 24013 24053
FT DOMAIN 24053 24093
FT DOMAIN 24093 24133
FT DOMAIN 24133 24173
FT DOMAIN 24173 24213
FT DOMAIN 24213 24253
FT DOMAIN 24253 24293
FT DOMAIN 24293 24333
FT DOMAIN 24333 24373
FT DOMAIN 24373 24413
FT DOMAIN 24413 24453
FT DOMAIN 24453 24493
FT DOMAIN 24493 24533
FT DOMAIN 24533 24573
FT DOMAIN 24573 24613
FT DOMAIN 24613 24653
FT DOMAIN 24653 24693
FT DOMAIN 24693 24733
FT DOMAIN 24733 24773
FT DOMAIN 24773 24813
FT DOMAIN 24813 24853
FT DOMAIN 24853 24893
FT DOMAIN 24893 24933
FT DOMAIN 24933 24973
FT DOMAIN 24973 25013
FT DOMAIN 25013 25053
FT DOMAIN 25053 25093
FT DOMAIN 25093 25133
FT DOMAIN 25133 25173
FT DOMAIN 25173 25213
FT DOMAIN 25213 25253
FT DOMAIN 25253 25293
FT DOMAIN 25293 25333
FT DOMAIN 25333 25373
FT DOMAIN 25373 25413
FT DOMAIN 25413 25453
FT DOMAIN 25453 25493
FT DOMAIN 25493 25533
FT DOMAIN 25533 25573
FT DOMAIN 25573 25613
FT DOMAIN 25613 25653
FT DOMAIN 25653 25693
FT DOMAIN 25693 25733
FT DOMAIN 25733 25773
FT DOMAIN 25773 25813
FT DOMAIN 25813 25853
FT DOMAIN 25853 25893
FT DOMAIN 25893 25933
FT DOMAIN 25933 25973
FT DOMAIN 25973 26013
FT DOMAIN 26013 26053
FT DOMAIN 26053 26093
FT DOMAIN 26093 26133
FT DOMAIN 26133 26173
FT DOMAIN 26173 26213
FT DOMAIN 26213 26253
FT DOMAIN 26253 26293
FT DOMAIN 26293 26333
FT DOMAIN 26333 26373
FT DOMAIN 26373 26413
FT DOMAIN 26413 26453
FT DOMAIN 26453 26493
FT DOMAIN 26493 26533
FT DOMAIN 26533 26573
FT DOMAIN 26573 26613
FT DOMAIN 26613 26653
FT DOMAIN 26653 26693
FT DOMAIN 26693 26733
FT DOMAIN 26733 26773
FT DOMAIN 26773 26813
FT DOMAIN 26813 26853
FT DOMAIN 26853 26893
FT DOMAIN 26893 26933
FT DOMAIN 26933 26973
FT DOMAIN 26973 27013
FT DOMAIN 27013 27053
FT DOMAIN 27053 27093
FT DOMAIN 27093 27133
FT DOMAIN 27133 27173
FT DOMAIN 27173 27213
FT DOMAIN 27213 27253
FT DOMAIN 27253 27293
FT DOMAIN 27293 27333
FT DOMAIN 27333 27373
FT DOMAIN 27373 27413
FT DOMAIN 27413 27453
FT DOMAIN 27453 27493
FT DOMAIN 27493 27533
FT DOMAIN 27533 27573
FT DOMAIN 27573 27613
FT DOMAIN 27613 27653
FT DOMAIN 27653 27693
FT DOMAIN 27693 27733
FT DOMAIN 27733 27773
FT DOMAIN 27773 27813
FT DOMAIN 27813 27853
FT DOMAIN 27853 27893
FT DOMAIN 27893 27933
FT DOMAIN 27933 27973
FT DOMAIN 27973 28013
FT DOMAIN 28013 28053
FT DOMAIN 28053 28093
FT DOMAIN 28093 28133
FT DOMAIN 28133 28173
FT DOMAIN 28173 28213
FT DOMAIN 28213 28253
FT DOMAIN 28253 28293
FT DOMAIN 28293 28333
FT DOMAIN 28333 28373
FT DOMAIN 28373 28413
FT DOMAIN 28413 28453
FT DOMAIN 28453 28493
FT DOMAIN 28493 28533
FT DOMAIN 28533 28573
FT DOMAIN 28573 28613
FT DOMAIN 28613 28653
FT DOMAIN 28653 28693
FT DOMAIN 28693 28733
FT DOMAIN 28733 28773
FT DOMAIN 28773 28813
FT DOMAIN 28813 28853
FT DOMAIN 28853 28893
FT DOMAIN 28893 28933
FT DOMAIN 28933 28973
FT DOMAIN 28973 29013
FT DOMAIN 29013 29053
FT DOMAIN 29053 29093
FT DOMAIN 29093 29133
FT DOMAIN 29133 29173
FT DOMAIN 29173 29213
FT DOMAIN 29213 29253
FT DOMAIN 29253 29293
FT DOMAIN 29293 29333
FT DOMAIN 29333 29373
FT DOMAIN 29373 29413
FT DOMAIN 29413 29453
FT DOMAIN 29453 29493
FT DOMAIN 29493 29533
FT DOMAIN 29533 29573
FT DOMAIN 29573 2961
```



FT	DOMAIN	2019	2060	EGF-LIKE 8.	
FT	DOMAIN	2108	2156	LDL-RECEPTOR CLASS B 20.	
FT	DOMAIN	2158	2201	LDL-RECEPTOR CLASS B 21.	
FT	DOMAIN	2203	2245	LDL-RECEPTOR CLASS B 22.	
FT	DOMAIN	2247	2289	LDL-RECEPTOR CLASS B 23.	
FT	DOMAIN	2291	2332	LDL-RECEPTOR CLASS B 24.	
FT	DOMAIN	2343	2384	EGF-LIKE 9.	
FT	DOMAIN	2432	2477	LDL-RECEPTOR CLASS B 25.	
FT	DOMAIN	2479	2518	LDL-RECEPTOR CLASS B 26.	
FT	DOMAIN	2520	2562	LDL-RECEPTOR CLASS B 27.	
FT	DOMAIN	2564	2604	LDL-RECEPTOR CLASS B 28.	
FT	DOMAIN	2605	2647	LDL-RECEPTOR CLASS B 29.	
FT	DOMAIN	2652	2694	EGF-LIKE 10.	
FT	DOMAIN	2699	2739	LDL-RECEPTOR CLASS A 16.	
FT	DOMAIN	2740	2778	LDL-RECEPTOR CLASS A 17.	
FT	DOMAIN	2779	2820	LDL-RECEPTOR CLASS A 18.	
FT	DOMAIN	2821	2862	LDL-RECEPTOR CLASS A 19.	
FT	DOMAIN	2863	2903	LDL-RECEPTOR CLASS A 20.	
FT	DOMAIN	2906	2947	LDL-RECEPTOR CLASS A 21.	
FT	DOMAIN	2948	2992	LDL-RECEPTOR CLASS A 22.	
FT	DOMAIN	2993	3031	LDL-RECEPTOR CLASS A 23.	
FT	DOMAIN	3032	3072	LDL-RECEPTOR CLASS A 24.	
FT	DOMAIN	3075	3112	LDL-RECEPTOR CLASS A 25.	
FT	DOMAIN	3113	3153	EGF-LIKE 11.	
FT	DOMAIN	3154	3194	EGF-LIKE 12.	
FT	DOMAIN	3241	3282	LDL-RECEPTOR CLASS B 30.	
FT	DOMAIN	3284	3333	LDL-RECEPTOR CLASS B 31.	
FT	DOMAIN	3335	3377	LDL-RECEPTOR CLASS B 32.	
FT	DOMAIN	3379	3420	LDL-RECEPTOR CLASS B 33.	
FT	DOMAIN	3421	3461	LDL-RECEPTOR CLASS B 34.	
FT	DOMAIN	3462	3511	EGF-LIKE 13.	
FT	DOMAIN	3512	3552	LDL-RECEPTOR CLASS A 26.	
FT	DOMAIN	3553	3593	LDL-RECEPTOR CLASS A 27.	
FT	DOMAIN	3594	3634	LDL-RECEPTOR CLASS A 28.	
FT	DOMAIN	3635	3675	LDL-RECEPTOR CLASS A 29.	
FT	DOMAIN	3678	3718	LDL-RECEPTOR CLASS A 30.	
FT	DOMAIN	3719	3758	LDL-RECEPTOR CLASS A 31.	
FT	DOMAIN	3759	3797	LDL-RECEPTOR CLASS A 32.	
FT	DOMAIN	3798	3836	LDL-RECEPTOR CLASS A 33.	
FT	DOMAIN	3842	3882	LDL-RECEPTOR CLASS A 34.	
FT	DOMAIN	3883	3924	LDL-RECEPTOR CLASS A 35.	
FT	DOMAIN	3928	3966	LDL-RECEPTOR CLASS A 36.	
FT	DOMAIN	3968	4008	EGF-LIKE 14.	
FT	DOMAIN	4009	4050	EGF-LIKE 15.	
FT	DOMAIN	4156	4197	LDL-RECEPTOR CLASS B 35.	
FT	DOMAIN	4199	4241	LDL-RECEPTOR CLASS B 36.	
FT	DOMAIN	4244	4284	LDL-RECEPTOR CLASS B 37.	
FT	DOMAIN	4332	4370	EGF-LIKE 16.	
FT	DOMAIN	4379	4413	EGF-LIKE 17.	
FT	SITE	4454	4460	SH3-BINDING (POTENTIAL).	
FT	SITE	4457	4463	SH3-BINDING (POTENTIAL).	
FT	SITE	4606	4609	SH2-BINDING (POTENTIAL).	
FT	SITE	4619	4625	SH3-BINDING (POTENTIAL).	
FT	SITE	4624	4630	SH3-BINDING (POTENTIAL).	
FT	SITE	1743	1745	CELL ATTACHMENT SITE (POTENTIAL).	
FT	SITE	4522	4527	ENDOCYTOSIS SIGNAL (POTENTIAL).	
FT	SITE	4601	4606	ENDOCYTOSIS SIGNAL (POTENTIAL).	
FT	SITE	28	40	BY SIMILARITY.	
FT	DISULFID	35	53	BY SIMILARITY.	
FT	DISULFID	47	62	BY SIMILARITY.	
FT	DISULFID	67	80	BY SIMILARITY.	
FT	DISULFID	74	93	BY SIMILARITY.	
FT	DISULFID	87	103	BY SIMILARITY.	
FT	DISULFID	108	120	BY SIMILARITY.	
Query Match					39.98; Score 202; DB 1; Length 4660;
Best Local Similarity					45.3%; Pred. No. 1.2e-11;
Matches					39; Conservative 11; Mismatches 32; Indels 4; Gaps 3;
Qy	1	KTCSFKQFACRQDITICISKGRCDGDCPDGSDAPEICPQSKAQRCQNEHNCIGTEL	60		
Db	1270	KTCSPTFLC-DNGNCIYKAWICDGDNDRCRMSDE--KDCP-QPFCFSTQCCPGYST	1325		

Qy	61	CVPMRLCNGVQCMDCGSDGPHCRE	86		
Db	1326	CINLSALCDGVDFCPNGTDSPLCNQ	1351		
RESULT 5					
YL_DROME	YL_DROME	STANDARD;	PRT;	1984	AA.
AC	P98163;				
DT	01-OCT-1996 (Rel. 34, Created)				
DT	01-OCT-1996 (Rel. 34, Last sequence update)				
DT	10-OCT-2003 (Rel. 42, Last annotation update)				
DE	Putative vitellogenin receptor precursor (Yolkless protein) (YL).				
GN	YL.				
OS	Drosophila melanogaster (Fruit fly).				
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
OC	Ephydroidea; Drosophilidae; Drosophila.				
OX	NCBI_taxID=7227;				
RN	11				
RP	SEQUENCE FROM N.A.				
RX	TISSUE=Ovary;				
RC	MEDLINE=95183490; PubMed=7878005;				
RA	Schonbaum C.P., Lee S., Mahowald A.P.;				
RT	"The Drosophila yolkless gene encodes a vitellogenin receptor				
RT	belonging to the low density lipoprotein receptor superfamily."				
RL	Proc Natl Acad Sci U.S.A. 92:1485-1489(1995).				
CC	-!- FUNCTION: Involved in uptake of vitellogenin by endocytosis.				
CC	-!- TISSUE SPECIFICITY: Ovary.				
CC	-!- SIMILARITY: Contains 13 LDL-receptor class A domains.				
CC	-!- SIMILARITY: Contains 7 EGF-like domains.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC	or send an email to license@sb-sib.ch).				
CC	-----				
DR	EMBL; U13637; AAB60217.1; -.				
DR	PIR; T13171; T13171.				
DR	HSSP; P01130; 1AJJ.				
DR	FlyBase; F5gn004649; YL.				
DR	GO; GO:007292; P:female gamete generation; NAS.				
DR	InterPro; IPR001152; Asx hydroxyl_S.				
DR	InterPro; IPR001881; EGF_Ca.				
DR	InterPro; IPR006209; EGF_like.				
DR	InterPro; IPR002172; Ldl_receptor_A.				
DR	InterPro; IPR000033; Ldl_receptor_rep.				
DR	Pfam; PF00008; EGF; 5.				
DR	Pfam; PF00057; ldl_recept_a; 13.				
DR	Pfam; PF00058; ldl_recept_b; 6.				
DR	PRINTS; PR00361; LDLRECEPTOR.				
DR	SMART; SM00199; EGF_CA; 2.				
DR	SMART; SM00192; LDLa; 13.				
DR	SMART; SM00135; LY; 10.				
DR	PROSITE; PS00010; ASX_HYDROXYL; 2.				
DR	PROSITE; PS01186; EGF_2; 3.				
DR	PROSITE; PS00026; EGF_3; 3.				
DR	PROSITE; PS01187; EGF_CA; 2.				
DR	PROSITE; PS01209; LDLRA_1; 11.				
DR	PROSITE; PS00068; LDLRA_2; 13.				
KW	Glycoprotein; Signal; Transmembrane; EGF-like domain; Receptor;				
KW	Repeat; Endocytosis.				
FT	SIGNAL	1	?		POTENTIAL.
FT	CHAIN	?	1984		POTENTIAL.
FT	DOMAIN	?	1800		EXTRACELLULAR (POTENTIAL).
FT	DOMAIN	1801	1821		POTENTIAL.
FT	TRANSMEM	1822	1984		CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	88	126		LDL-RECEPTOR CLASS A 1.
FT	DOMAIN	127	167		LDL-RECEPTOR CLASS A 2.
FT	DOMAIN	182	222		LDL-RECEPTOR CLASS A 3.

95.2

```

FT DOMAIN 225 264 LDL-RECEPTOR CLASS A 4.
FT DOMAIN 264 306 LDL-RECEPTOR CLASS A 5.
FT DOMAIN 306 347 EGF-LIKE 1.
FT DOMAIN 348 388 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 660 701 EGF-LIKE 3.
FT DOMAIN 984 1026 EGF-LIKE 4.
FT DOMAIN 1029 1064 LDL-RECEPTOR CLASS A 6.
FT DOMAIN 1072 1111 LDL-RECEPTOR CLASS A 7.
FT DOMAIN 1116 1154 LDL-RECEPTOR CLASS A 8.
FT DOMAIN 1156 1195 LDL-RECEPTOR CLASS A 9.
FT DOMAIN 1234 1273 LDL-RECEPTOR CLASS A 10.
FT DOMAIN 1281 1320 LDL-RECEPTOR CLASS A 11.
FT DOMAIN 1320 1359 LDL-RECEPTOR CLASS A 12.
FT DOMAIN 1338 1377 LDL-RECEPTOR CLASS A 13.
FT DOMAIN 1375 1417 EGF-LIKE 5.
FT DOMAIN 1418 1457 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1734 1770 EGF-LIKE 7.
FT SITE 1837 1877 CRITICAL FOR ENDOCYTOSIS (POTENTIAL).
FT SITE 1878 1918 CRITICAL FOR ENDOCYTOSIS (POTENTIAL).
FT SITE 1892 1932 CRITICAL FOR ENDOCYTOSIS (POTENTIAL).
FT SITE 90 102 BY SIMILARITY.
FT DISULFID 97 115 BY SIMILARITY.
FT DISULFID 109 124 BY SIMILARITY.
FT DISULFID 129 144 BY SIMILARITY.
FT DISULFID 137 157 BY SIMILARITY.
FT DISULFID 151 166 BY SIMILARITY.
FT DISULFID 184 197 BY SIMILARITY.
FT DISULFID 191 210 BY SIMILARITY.
FT DISULFID 204 220 BY SIMILARITY.
FT DISULFID 227 239 BY SIMILARITY.
FT DISULFID 234 253 BY SIMILARITY.
FT DISULFID 247 262 BY SIMILARITY.
FT DISULFID 266 281 BY SIMILARITY.
FT DISULFID 275 294 BY SIMILARITY.
FT DISULFID 298 304 BY SIMILARITY.
FT DISULFID 310 321 BY SIMILARITY.
FT DISULFID 315 331 BY SIMILARITY.
FT DISULFID 333 346 BY SIMILARITY.
FT DISULFID 352 363 BY SIMILARITY.
FT DISULFID 359 372 BY SIMILARITY.
FT DISULFID 374 387 BY SIMILARITY.
FT DISULFID 664 673 BY SIMILARITY.
FT DISULFID 669 685 BY SIMILARITY.
FT DISULFID 687 700 BY SIMILARITY.
FT DISULFID 988 999 BY SIMILARITY.
FT DISULFID 995 1009 BY SIMILARITY.
FT DISULFID 1011 1025 BY SIMILARITY.
FT DISULFID 1031 1040 BY SIMILARITY.
FT DISULFID 1035 1053 BY SIMILARITY.
FT DISULFID 1047 1062 BY SIMILARITY.
FT DISULFID 1074 1087 BY SIMILARITY.
FT DISULFID 1081 1100 BY SIMILARITY.
FT DISULFID 1094 1109 BY SIMILARITY.
FT DISULFID 1118 1130 BY SIMILARITY.
FT DISULFID 1125 1143 BY SIMILARITY.
FT DISULFID 1137 1152 BY SIMILARITY.
FT DISULFID 1158 1170 BY SIMILARITY.
FT DISULFID 1165 1183 BY SIMILARITY.
FT DISULFID 1177 1193 BY SIMILARITY.
FT DISULFID 1198 1210 BY SIMILARITY.
FT DISULFID 1217 1232 BY SIMILARITY.
FT DISULFID 1243 1257 BY SIMILARITY.
FT DISULFID 1250 1270 BY SIMILARITY.
FT DISULFID 1264 1279 BY SIMILARITY.
FT DISULFID 1283 1296 BY SIMILARITY.
FT DISULFID 1290 1309 BY SIMILARITY.
FT DISULFID 1303 1318 BY SIMILARITY.
FT DISULFID 1340 1352 BY SIMILARITY.
FT DISULFID 1347 1365 BY SIMILARITY.
FT DISULFID 1359 1375 BY SIMILARITY.
FT DISULFID 1379 1392 BY SIMILARITY.
FT DISULFID 1388 1401 BY SIMILARITY.

Query Match 39.0%; Score 197.5; DB 1; Length 1984;
Best Local Similarity 42.6%; Pred. No. 1.5e-11;
Matches 40; Conservative 13; Mismatches 28; Indels 13; Gaps 3;

Qy 3 CSPKQFACRDQITCISKWRCDCGERDCPDGDS-----APEICPQSKAQR-CQP 50
Db 1283 CSIVFKCRSGRECIRREFCDQGDGSDLSCELEKHGHNQIQIWFSTSSRSCP 1342
Qy 51 NEHNCIGTELCPVMSRLCNQVQCDMGDSDEGPHC 84
Db 1343 HLFDCDQGE-CVDSLRVCNNFPDCTNGHDEGPKC 1375

RESULT 6
LDL1_XENLA
ID LDL1_XENLA STANDARD; PRT; 909 AA.
AC Q99087;
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Low-density lipoprotein receptor 1 precursor (LDL receptor 1).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver, and Oocyte;
RA MEDLINE=5124815; PubMed=1709931;
RX Menta K.D., Chen W.J., Goldstein J.L., Brown M.S.;
RT "The low density lipoprotein receptor in Xenopus laevis. I. Five
RT domains that resemble the human receptor.";
RL J. Biol. Chem. 266:10406-10414(1991).
CC -!- FUNCTION: Binds LDL, the major cholesterol-carrying lipoprotein
CC of plasma, and transports it into cells by endocytosis. In order
CC to be internalized, the receptor-ligand complexes must first
CC cluster into clathrin-coated pits.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: Contains 7 LDL-receptor class A domains.
CC -!- SIMILARITY: Contains 6 LDL-receptor class B domains.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC EMBL; M62976; AAA49897.1; -.
DR HSSP; P01130; 1AJJ.
DR InterPro; IPR000152; Asx hydroxyl_s.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR004209; EGF_like.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000033; LDL_receptor_rep.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00057; ldl_recept_a; 7.
DR Pfam; PF00058; ldl_recept_b; 5.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SMO0179; EGF_CA; 1.
DR SMART; SMO0192; LDRA; 7.
DR SMART; SMO0135; LY; 5.
DR PROSITE; PS00010; ASX_HYDROXYL; 2.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS01209; LDLRA_1; 7.
DR PROSITE; PS00068; LDLRA_2; 7.
KW Glycoprotein; LDL; Cholesterol metabolism; Lipid transport;
KW Endocytosis; Coated pits; Transmembrane; Receptor; Signal;
KW EGF-like domain; Repeat.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 809 LOW-DENSITY LIPOPROTEIN RECEPTOR 1.
FT DOMAIN 22 836 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 837 858 POTENTIAL.
FT DOMAIN 859 909 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 25 65 LDL-RECEPTOR CLASS A 1.
FT DOMAIN 66 106 LDL-RECEPTOR CLASS A 2.
FT DOMAIN 107 145 LDL-RECEPTOR CLASS A 3.
FT DOMAIN 146 185 LDL-RECEPTOR CLASS A 4.
FT DOMAIN 193 231 LDL-RECEPTOR CLASS A 5.
FT DOMAIN 232 270 LDL-RECEPTOR CLASS A 6.
FT DOMAIN 272 311 LDL-RECEPTOR CLASS A 7.
FT DOMAIN 312 351 EGF-LIKE 1.
FT DOMAIN 352 391 EGF-LIKE 2.
FT REPEAT 395 436 LDL-RECEPTOR CLASS B 1.
FT REPEAT 437 483 LDL-RECEPTOR CLASS B 2.
FT REPEAT 484 526 LDL-RECEPTOR CLASS B 3.
FT REPEAT 527 570 LDL-RECEPTOR CLASS B 4.
FT REPEAT 571 613 LDL-RECEPTOR CLASS B 5.
FT REPEAT 614 655 LDL-RECEPTOR CLASS B 6.
FT DOMAIN 661 710 EGF-LIKE 3.
FT DOMAIN 717 813 CLUSTERED O-LINKED OLIGOSACCHARIDES.
FT SITE 871 876 ENDOCYTOSIS SIGNAL (POTENTIAL).
FT DISULFID 27 39 BY SIMILARITY.
FT DISULFID 34 52 BY SIMILARITY.
FT DISULFID 46 63 BY SIMILARITY.
FT DISULFID 68 82 BY SIMILARITY.
FT DISULFID 75 95 BY SIMILARITY.
FT DISULFID 89 104 BY SIMILARITY.
FT DISULFID 109 121 BY SIMILARITY.
FT DISULFID 116 134 BY SIMILARITY.
FT DISULFID 128 143 BY SIMILARITY.
FT DISULFID 148 161 BY SIMILARITY.
FT DISULFID 155 174 BY SIMILARITY.
FT DISULFID 168 183 BY SIMILARITY.
FT DISULFID 195 207 BY SIMILARITY.
FT DISULFID 202 220 BY SIMILARITY.
FT DISULFID 214 229 BY SIMILARITY.
FT DISULFID 234 246 BY SIMILARITY.
FT DISULFID 241 259 BY SIMILARITY.
FT DISULFID 253 268 BY SIMILARITY.
FT DISULFID 274 287 BY SIMILARITY.
FT DISULFID 282 300 BY SIMILARITY.
FT DISULFID 294 311 BY SIMILARITY.
FT DISULFID 316 327 BY SIMILARITY.
FT DISULFID 323 336 BY SIMILARITY.
FT DISULFID 338 350 BY SIMILARITY.
FT DISULFID 356 366 BY SIMILARITY.

FT DISULFID 362 375 BY SIMILARITY.
FT DISULFID 377 390 BY SIMILARITY.
FT DISULFID 665 679 BY SIMILARITY.
FT DISULFID 675 694 BY SIMILARITY.
FT DISULFID 696 709 BY SIMILARITY.
FT CARBOHYD 97 97 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 270 270 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 909 AA; 101295 MW; 6ED41F5402A16371 CRC64;

Query Match 38.0%; Score 192.5; DB 1; Length 909;
Best Local Similarity 45.7%; Pred. No. 2.3e-11;
Matches 42; Conservative 9; Mismatches 30; Indels 11; Gaps 4;

QY 2 TCSPKQAFACRDQITCISKWRCDDGERDCPDGSDGAPEIC----POSKAQRQCPNEHNCIG 57
Db 147 TCNPAMFOCKDKGICIPKLWACDGDPPDCEDGSDSDE--EHCEGREPIKTKPCSPLEFHC-G 203
QY 58 TELCVPMSELNGLVQDCMDGSDSDE-----GPHCR 85
Db 204 SGECHMSWKCDGGDFCDKSKSDEKDCVKTCTCR 235

RESULT 7
LDLR RAT
ID LDLR RAT STANDARD; PRT; 879 AA.
AC P33952;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Low-density lipoprotein receptor precursor (LDL receptor).
GN LDLR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN:Sprague-Dawley;
RX MEDLINE=89160263; PubMed=2922268;
RA Lee L.Y., Mohler W.A., Schafer B.D., Freudenberger J.S.,
RA Byrne-Connolly N., Eager K.B., Mosley S.T., Leighton J.K.,
RA Thrift R.N., Davis R.A., Tanaka R.D.;
RT "Nucleotide sequence of the rat low density lipoprotein receptor
RT cDNA."
RL Nucleic Acids Res. 17:1259-1260(1989).
CC -!- FUNCTION: Binds LDL, the major cholesterol-carrying lipoprotein of
CC plasma, and transports it into cells by endocytosis. In order to
CC be internalized, the receptor-ligand complexes must first cluster
CC into clathrin-coated pits.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: Contains 7 LDL-receptor class A domains.
CC -!- SIMILARITY: Contains 6 LDL-receptor class B domains.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X13722; CAA32001.1; -.
DR PIR; S03430; QRRULD.
DR HSSP; P01130; 1AJJ.
DR InterPro; IPR000152; Asx hydroxyl_s.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000033; Ldl_receptor_rep.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF00057; ldl_recept_a; 7.

```





CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC

DR EMBL; U08110; CAA69325.1; -  
DR HSSP; P01130; LAJJ  
DR Genew; HGNC:11185; SORL1.  
DR MLM; 602005; -  
DR GO; GO:0005887; C.integral to plasma membrane; TAS.  
DR GO; GO:0015029; F.internationalization receptor activity; TAS.  
DR GO; GO:0004888; F.transmembrane receptor activity; TAS.  
DR GO; GO:0006898; P.receptor mediated endocytosis; TAS.  
DR InterPro; IPR006209; EGF like.  
DR InterPro; IPR008957; FN III-like.  
DR InterPro; IPR003961; FN III.  
DR InterPro; IPR002860; GH\_BNR.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR002172; LDL receptor A.  
DR InterPro; IPR000033; LDL receptor\_rep.  
DR Pfam; PF02012; BNR; 5.  
DR Pfam; PF00041; fn3; 5.  
DR Pfam; PF00057; ldl\_recept\_a; 11.  
DR Pfam; PF00058; ldl\_recept\_b; 4.  
DR PRINTS; PR00261; LDLRECEPTOR.  
DR SMART; SM00181; EGF; 1.  
DR SMART; SM00060; FN3; 6.  
DR SMART; SM00192; LDLa; 11.  
DR SMART; SM00135; LX; 5.  
DR SMART; SM00602; VPS10; 1.  
DR PROSITE; PS01186; EGF\_2; 1.  
DR PROSITE; PS01209; LDLRA\_1; 10.  
DR PROSITE; PS00068; LDLRA\_2; 11.  
KW Endocytosis; Receptor; Transmembrane; EGF-like domain; Repeat; Glycoprotein; LDL; Lipid transport; Cholesterol metabolism; Signal.  
FT SIGNAL 1 28 POTENTIAL.  
FT PROPEP 29 81 REMOVED IN MATURE PROTEIN.  
FT CHAIN 82 2214 SORTLIN-RELATED RECEPTOR.  
FT DOMAIN 82 2137 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 2138 2158 POTENTIAL.  
FT DOMAIN 2159 2214 CYTOPLASMIC (POTENTIAL).  
FT REPEAT 136 147 BNR 1.  
FT REPEAT 232 243 BNR 2.  
FT REPEAT 441 452 BNR 3.  
FT REPEAT 521 532 BNR 4.  
FT REPEAT 562 573 BNR 5.  
FT DOMAIN 803 977 5 X APPROXIMATE YWTD REPEATS.  
FT REPEAT 803 806 1.  
FT REPEAT 847 850 2.  
FT REPEAT 891 894 3.  
FT REPEAT 934 937 4.  
FT REPEAT 974 977 5.  
FT DOMAIN 1026 1072 EGF-LIKE.  
FT DOMAIN 1076 1114 LDL-RECEPTOR CLASS A 1.  
FT DOMAIN 1115 1155 LDL-RECEPTOR CLASS A 2.  
FT DOMAIN 1156 1194 LDL-RECEPTOR CLASS A 3.  
FT DOMAIN 1197 1237 LDL-RECEPTOR CLASS A 4.  
FT DOMAIN 1237 1273 LDL-RECEPTOR CLASS A 5.  
FT DOMAIN 1273 1317 LDL-RECEPTOR CLASS A 6.  
FT DOMAIN 1323 1361 LDL-RECEPTOR CLASS A 7.  
FT DOMAIN 1366 1405 LDL-RECEPTOR CLASS A 8.  
FT DOMAIN 1417 1455 LDL-RECEPTOR CLASS A 9.  
FT DOMAIN 1469 1508 LDL-RECEPTOR CLASS A 10.  
FT DOMAIN 1512 1551 LDL-RECEPTOR CLASS A 11.  
FT DOMAIN 1556 1645 FIBRONECTIN TYPE-III 1.  
FT DOMAIN 1653 1742 FIBRONECTIN TYPE-III 2.  
FT DOMAIN 1749 1837 FIBRONECTIN TYPE-III 3.  
FT DOMAIN 1842 1927 FIBRONECTIN TYPE-III 4.  
FT DOMAIN 1933 2024 FIBRONECTIN TYPE-III 5.  
FT DOMAIN 2025 2115 FIBRONECTIN TYPE-III 6.  
FT SITE 63 CELL ATTACHMENT SITE (POTENTIAL).  
FT SITE 65 ENDOCYTOSIS SIGNAL (POTENTIAL).  
FT DISULFID 1078 1090 BY SIMILARITY.

FT DISULFID 1085 1103 BY SIMILARITY.  
FT DISULFID 1097 1112 BY SIMILARITY.  
FT DISULFID 1117 1131 BY SIMILARITY.  
FT DISULFID 1125 1144 BY SIMILARITY.  
FT DISULFID 1138 1158 BY SIMILARITY.  
FT DISULFID 1158 1170 BY SIMILARITY.  
FT DISULFID 1165 1183 BY SIMILARITY.  
FT DISULFID 1177 1192 BY SIMILARITY.  
FT DISULFID 1199 1211 BY SIMILARITY.  
FT DISULFID 1206 1224 BY SIMILARITY.  
FT DISULFID 1218 1235 BY SIMILARITY.  
FT DISULFID 1239 1249 BY SIMILARITY.  
FT DISULFID 1244 1262 BY SIMILARITY.  
FT DISULFID 1256 1315 BY SIMILARITY.  
FT DISULFID 1325 1337 BY SIMILARITY.  
FT DISULFID 1332 1350 BY SIMILARITY.  
FT DISULFID 1344 1359 BY SIMILARITY.  
FT DISULFID 1368 1381 BY SIMILARITY.  
FT DISULFID 1376 1394 BY SIMILARITY.  
FT DISULFID 1388 1403 BY SIMILARITY.  
FT DISULFID 1426 1444 BY SIMILARITY.  
FT DISULFID 1438 1453 BY SIMILARITY.  
FT CARBOHYD 99 99 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 158 158 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 368 368 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 430 430 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 674 674 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 818 818 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 871 871 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1035 1035 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1068 1068 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1164 1164 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1191 1191 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1246 1246 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1367 1367 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1458 1458 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1608 1608 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1706 1706 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1733 1733 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1809 1809 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1854 1854 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1894 1894 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1986 1986 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 2010 2010 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 2054 2054 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 2069 2069 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 2076 2076 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 2092 2092 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 2214 AA; 248439 MW; 83F14BFBEB23C0C CRC64;  
Query Match 37.18; Score 187.5; DB 1; Length 2214;  
Best Local Similarity 47.6%; Pred. No. 1; ee-10;  
Matches 40; Conservative 9; Mismatches 24; Indels 11; Gaps 6;  
QY 2 TCSPKQFACRDQITCISKGRCDGERDCPDGSDPAEICPOS-KAQCQNEHNCIGTEL 60  
Db 1198 TCEASNFQCRNG-HCIPQRWACDGTDCQDGSDEDPVNCCKGKNGFC-PN-----GT-- 1248  
QY 61 CVFNSRLCNGVQDCMGDSDEGPHC 84  
Db 1249 CIPSSKCHDGLRDCSDGSDE-QHC 1271  
RESULT 10  
UN52 CAEL STANDARD; PRT; 3375 AA.  
ID UN52 CAEL Q08561; O18263; Q9XTD2; Q9XTI5;  
AC Q08561; O18263; Q9XTD2; Q9XTI5;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Basement membrane proteoglycan precursor (Perlecan homolog)  
DE (Uncoordinated protein 52).

GN UNC-52 OR ZC101.2.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM A), AND FUNCTION.  
RX MEDLINE=9339574; PubMed=6393416;  
RA Regalski T.M., Williams B.D., Mullen G.P., Moerman D.G.;  
RT "Products of the unc-52 gene in Caenorhabditis elegans are homologous  
RT to the core protein of the mammalian basement membrane heparan  
RT sulfate proteoglycan";  
RL Genes Dev. 7:1471-1484 (1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Percy C.M., Baynes C.;  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP REVISIONS, AND ALTERNATIVE SPLICING.  
RA Durbin R.;  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Probable role in myofibrillar assembly and/or attachment  
CC of the myofibrillar lattice to the cell membrane. May be an  
CC extracellular anchor for integrin receptors in muscle.  
CC -!- SUBCELLULAR LOCATION: Extracellular matrix.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=4;  
CC Name=a;  
CC IsoId=Q06561-1; Sequence=Displayed;  
CC Note=No experimental confirmation available;  
CC Name=a;  
CC IsoId=Q06561-2; Sequence=VSP\_007195, VSP\_007196;  
CC Name=b;  
CC IsoId=Q06561-3; Sequence=VSP\_007191, VSP\_007192;  
CC Note=No experimental confirmation available;  
CC Name=c;  
CC IsoId=Q06561-4; Sequence=VSP\_007193, VSP\_007194, VSP\_007195,  
CC VSP\_007196;  
CC Note=No experimental confirmation available;  
CC -!- TISSUE SPECIFICITY: Found in the basement membrane of all  
CC contractile tissues. It is concentrated over muscle dense bodies  
CC and M-lines which are associated with beta-integrin.  
CC -!- DEVELOPMENTAL STAGE: Synthesized early in embryogenesis.  
CC -!- SIMILARITY: Contains 3 LDL-receptor class A domains.  
CC -!- SIMILARITY: Contains 16 immunoglobulin-like C2-type domains.  
CC -!- SIMILARITY: Contains 7 laminin G-like domains.  
CC -!- SIMILARITY: Contains 3 laminin G-like domains.  
CC -!- SIMILARITY: Contains 2 laminin IV domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; L13458; AAA28156.1; -  
CC EMBL; Z93375; CAB07567.1; -  
CC EMBL; Z93395; CAB07567.1; JOINED.  
CC EMBL; Z93375; CAB07568.1; -  
CC EMBL; Z93395; CAB07568.1; JOINED.  
CC EMBL; Z93375; CAB07569.1; -  
CC EMBL; Z93395; CAB07569.1; JOINED.  
CC EMBL; Z93395; CAB07704.1; -  
CC EMBL; Z93395; CAB07704.1; JOINED.  
CC EMBL; Z93395; CAB07706.1; -  
CC EMBL; Z93375; CAB07706.1; JOINED.  
CC EMBL; Z93395; CAB07707.1; -  
CC EMBL; Z93375; CAB07707.1; JOINED.  
CC EMBL; Z93395; CAB07708.1; -  
CC EMBL; Z93375; CAB07708.1; JOINED.  
CC HSSP; P01130; 1LDR.

DR WormPep; ZC101.2a; CE15028.  
DR WormPep; ZC101.2b; CE15030.  
DR WormPep; ZC101.2c; CE15034.  
DR WormPep; ZC101.2e; CE18424.  
DR GO; GO:0005578; C:extracellular matrix; IEP.  
DR GO; GO:0030239; P:myofibril assembly; IEP.  
DR InterPro; IPR008985; ConA\_like\_lec\_gl.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR006209; EGF\_like.  
DR InterPro; IPR006210; IEGF\_like.  
DR InterPro; IPR007110; IG-like.  
DR InterPro; IPR003599; IG.  
DR InterPro; IPR003598; IG\_c2.  
DR InterPro; IPR000034; Laminin\_B.  
DR InterPro; IPR002049; Laminin\_EGF.  
DR InterPro; IPR001791; Laminin\_G.  
DR InterPro; IPR002172; LDL\_receptor\_A.  
DR Pfam; PF00047; Ig; 16.  
DR Pfam; PF00052; laminin\_B; 2.  
DR Pfam; PF00053; laminin\_EGF; 5.  
DR Pfam; PF00057; ldl\_recept\_a; 3.  
DR PRINTS; PRO0261; LDLRECEPTOR.  
DR ProDom; PD003031; Laminin\_B; 2.  
DR SMART; SM0018; EGF; 6.  
DR SMART; SM00179; EGF\_CA; 2.  
DR SMART; SM00180; EGF\_Lam; 6.  
DR SMART; SM00408; IGC2; 17.  
DR SMART; SM00281; LamB; 1.  
DR SMART; SM00282; LamG; 3.  
DR SMART; SM00192; LDLA; 3.  
DR PROSITE; PS00022; EGF\_1; 7.  
DR PROSITE; PS01186; EGF\_2; 2.  
DR PROSITE; PS00026; EGF\_3; 3.  
DR PROSITE; PS00835; IG\_Like; 17.  
DR PROSITE; PS00025; LAM\_G\_DOMAIN; 3.  
DR PROSITE; PS01248; LAMININ\_TYPE\_EGF; 7.  
DR PROSITE; PS01209; LDLRA\_1; 3.  
DR PROSITE; PS00068; LDLRA\_2; 3.  
KW Proteoglycan; Glycoprotein; Signal; Immunoglobulin domain; Repeat;  
KW Basement membrane; Extracellular matrix; Alternative splicing;  
FT Laminin EGF-like domain.  
FT SIGNAL 1 22  
FT CHAIN 23 3375  
FT DOMAIN 45 130  
FT DOMAIN 148 184  
FT DOMAIN 189 225  
FT DOMAIN 232 269  
FT DOMAIN 271 355  
FT DOMAIN 384 431  
FT DOMAIN 432 441  
FT DOMAIN 442 633  
FT DOMAIN 634 666  
FT DOMAIN 674 720  
FT DOMAIN 721 730  
FT DOMAIN 731 921  
FT DOMAIN 922 954  
FT DOMAIN 955 1004  
FT DOMAIN 1011 1060  
FT DOMAIN 1061 1111  
FT DOMAIN 1126 1222  
FT DOMAIN 1226 1311  
FT DOMAIN 1319 1401  
FT DOMAIN 1410 1499  
FT DOMAIN 1503 1585  
FT DOMAIN 1588 1680  
FT DOMAIN 1690 1785  
FT DOMAIN 1793 1878  
FT DOMAIN 1886 1970  
FT DOMAIN 1973 2069  
FT DOMAIN 2073 2163  
FT DOMAIN 2173 2260  
FT DOMAIN 2263 2343  
POTENTIAL.  
BASEMENT MEMBRANE PROTEOGLYCAN.  
IG-LIKE C2-TYPE 1.  
LDL-RECEPTOR CLASS A 1.  
LDL-RECEPTOR CLASS A 2.  
LDL-RECEPTOR CLASS A 3.  
IG-LIKE C2-TYPE 2.  
LAMININ EGF-LIKE 1 (INCOMPLETE).  
LAMININ EGF-LIKE 2 (N-TERMINAL).  
LAMININ DOMAIN IV 1.  
LAMININ EGF-LIKE 2 (C-TERMINAL).  
LAMININ EGF-LIKE 3 (INCOMPLETE).  
LAMININ EGF-LIKE 4 (N-TERMINAL).  
LAMININ DOMAIN IV 2.  
LAMININ EGF-LIKE 4 (C-TERMINAL).  
LAMININ EGF-LIKE 5.  
LAMININ EGF-LIKE 6.  
LAMININ EGF-LIKE 7.  
IG-LIKE C2-TYPE 3.  
IG-LIKE C2-TYPE 4.  
IG-LIKE C2-TYPE 5.  
IG-LIKE C2-TYPE 6.  
IG-LIKE C2-TYPE 7.  
IG-LIKE C2-TYPE 8.  
IG-LIKE C2-TYPE 9.  
IG-LIKE C2-TYPE 10.  
IG-LIKE C2-TYPE 11.  
IG-LIKE C2-TYPE 12.  
IG-LIKE C2-TYPE 13.  
IG-LIKE C2-TYPE 14.  
IG-LIKE C2-TYPE 15.

FT	DOMAIN	2349	2435	IG-LIKE C2-TYPE 16.	OX	NCBI_TaxID=10090;
FT	DOMAIN	2446	2530	IG-LIKE C2-TYPE 17.	RN	[1]
FT	DOMAIN	2532	2713	LAMININ G-LIKE 1.	RP	SEQUENCE FROM N.A.
FT	DOMAIN	2793	2960	LAMININ G-LIKE 2.	RC	STRAIN=BALB/c; TISSUE=Heart;
FT	DOMAIN	2961	3093	GLU-RICH.	RX	MEDLINE=95010090; PubMed=7925422;
FT	DOMAIN	2972	3066	THR-RICH.	RA	Oka K., Ishimura-Oka K., Chu M.J., Sullivan M., Krushkal J.,
FT	DOMAIN	3180	3359	LAMININ G-LIKE 3.	RA	Li W.H., Chan L.;
FT	DISULFID	66	114	BY SIMILARITY.	RA	"Mouse very-low-density-lipoprotein receptor (VLDLR) cDNA cloning,
FT	DISULFID	149	161	BY SIMILARITY.	RT	tissue-specific expression and evolutionary relationship with the
FT	DISULFID	156	174	BY SIMILARITY.	RT	low-density-lipoprotein receptor.";
FT	DISULFID	168	183	BY SIMILARITY.	RL	Eur. J. Biochem. 224:975-982(1994).
FT	DISULFID	190	202	BY SIMILARITY.	RP	[2]
FT	DISULFID	197	215	BY SIMILARITY.	RN	SEQUENCE FROM N.A.
FT	DISULFID	209	224	BY SIMILARITY.	RC	TISSUE=Skeletal muscle;
FT	DISULFID	233	246	BY SIMILARITY.	RX	MEDLINE=94283285; PubMed=8013374;
FT	DISULFID	240	259	BY SIMILARITY.	RA	Gafvels M.E., Paavola L.G., Boyd C.O., Nolan P.M., Wittmaack F.,
FT	DISULFID	253	268	BY SIMILARITY.	RA	Chawla A., Lazar M.A., Bucan M., Angelin B.O., Strauss J.F.;
FT	DISULFID	955	964	BY SIMILARITY.	RT	"Cloning of a complementary deoxyribonucleic acid encoding the murine
FT	DISULFID	957	971	BY SIMILARITY.	RT	homolog of the very low density lipoprotein/apolipoprotein-E
FT	DISULFID	974	983	BY SIMILARITY.	RT	receptor: expression pattern and assignment of the gene to mouse
FT	DISULFID	986	1002	BY SIMILARITY.	RT	chromosome 19.";
FT	DISULFID	1011	1021	BY SIMILARITY.	RL	Endocrinology 135:387-394(1994).
FT	DISULFID	1013	1027	BY SIMILARITY.	RN	[3]
FT	DISULFID	1030	1039	BY SIMILARITY.	RP	SEQUENCE OF 204-262 FROM N.A.
FT	DISULFID	1042	1058	BY SIMILARITY.	RX	MEDLINE=95003355; PubMed=7919660;
FT	DISULFID	1061	1069	BY SIMILARITY.	RA	Neggert J.K., Wu J.L.;
FT	DISULFID	1063	1079	BY SIMILARITY.	RA	"The mouse very low density lipoprotein receptor (Vldlr) gene maps to
FT	DISULFID	1082	1091	BY SIMILARITY.	RT	chromosome 19.";
FT	DISULFID	1094	1109	BY SIMILARITY.	RL	Mamm. Genome 5:453-455(1994).
FT	DISULFID	1152	1200	BY SIMILARITY.	RN	[4]
FT	DISULFID	1338	1384	BY SIMILARITY.	RP	BINDING TO REELIN.
FT	DISULFID	1435	1481	BY SIMILARITY.	RX	MEDLINE=20036019; PubMed=10571241;
FT	DISULFID	1527	1573	BY SIMILARITY.	RA	Hiesberger T., Tromsdorff M., Howell B.W., Goffinet A.M., Mumby M.C.,
FT	DISULFID	1618	1663	BY SIMILARITY.	RA	Cooper J.A., Herz J.;
FT	DISULFID	1719	1767	BY SIMILARITY.	RT	"Direct binding of Reelin to VLDL receptor and ApoE receptor 2 induces
FT	DISULFID	1814	1861	BY SIMILARITY.	RT	tyrosine phosphorylation of disabled-1 and modulates tau
FT	DISULFID	1907	1954	BY SIMILARITY.	RL	phosphorylation.";
FT	DISULFID	1998	2053	BY SIMILARITY.	RL	Neuron 24:481-489(1999).
FT	DISULFID	2099	2147	BY SIMILARITY.	CC	-1- FUNCTION: Binds VLDL and transports it into cells by endocytosis.
FT	DISULFID	2195	2242	BY SIMILARITY.	CC	In order to be internalized, the receptor-ligand complexes must
FT	DISULFID	2284	2329	BY SIMILARITY.	CC	first cluster into clathrin-coated pits. Binding to Reelin induces
FT	DISULFID	2374	2420	BY SIMILARITY.	CC	tyrosine phosphorylation of Dab1 and modulation of tau
FT	DISULFID	2467	2514	BY SIMILARITY.	CC	phosphorylation.
FT	CARBOHYD	1422	1422	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	-1- SUBUNIT: Binds to the extracellular matrix protein Reelin.
FT	CARBOHYD	2476	2476	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.
FT	CARBOHYD	2950	2950	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	-1- TISSUE SPECIFICITY: Abundant in heart and muscle; less in kidney,
FT	CARBOHYD	3143	3143	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	brain, ovary, testis, lung and adipose tissue.
Qy	1	KTSPQKQFACDQITCISKWRCDCGERDCPDGSDPEICPSKQACOPNEHNCIGTEL 60			CC	-1- SIMILARITY: Contains 8 LDL-receptor class A domains.
Db	188	RTCEPNEFEKNNN-KCVQKWLDCGDDCDDNDEL-NCNAKPSDDCKPTBFQCHDRRQ 245			CC	-1- SIMILARITY: Contains 3 EGF-like domains.
Qy	61	CVPMRLCNGVQDCMGDSDE 80			CC	-----
Db	246	CVSSFFHCDGTNDCHGDSDE 265			CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
RESULT 11					CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
LDVR MOUSE					CC	the European Bioinformatics Institute. There are no restrictions on its
ID LDVR MOUSE					CC	use by non-profit institutions as long as its content is in no way
AC P98156; 064022; PRT; 873 AA.					CC	modified and this statement is not removed. Usage by and for commercial
DT 01-OCT-1996 (Rel. 34, Created)					CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>
DT 01-OCT-1996 (Rel. 34, Last sequence update)					CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).
DT 15-MAR-2004 (Rel. 43, Last annotation update)					CC	-----
DE Very low-density lipoprotein receptor precursor (VLDL receptor).					CC	EMBL; L33417; AAC37668.1; -;
GN VLDLR.					CC	EMBL; U06670; AAA59384.1; -;
OS Mus musculus (Mouse).					CC	EMBL; S73732; AAB32228.2; -;
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					CC	PIR; I48952; I48952.
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					CC	HSSP; P01130; IAJJ.
					CC	MGD; MGI:98935; Vldlr.
					CC	InterPro; IPR000152; Asx_hydroxyl_S.
					CC	InterPro; IPR001881; EGF_Ca.
					CC	InterPro; IPR006209; EGF_Like.
					CC	InterPro; IPR002172; LDL_receptor_A.
					CC	InterPro; IPR000033; Ldl_receptor_rep.
					CC	Pfam; PF00008; EGF; 2.
					CC	Pfam; PF00057; ldl_recept_a; 8.
					CC	Pfam; PF00058; ldl_recept_b; 5.



DR PRINTS; PRO0261; LDLRECEPTOR.  
DR SWART; SM00179; EGF CA; 2.  
DR SWART; SM00192; LDLA; 8.  
DR SWART; SM00195; LY; 5.  
DR PROSITE; PS00010; ASX HYDROXYL; 2.  
DR PROSITE; PS00022; EGF\_1; FALSE\_NEG.  
DR PROSITE; PS01186; EGF\_2; 3.  
DR PROSITE; PS00026; EGF\_3; 2.  
DR PROSITE; PS01187; EGF\_CA; 1.  
DR PROSITE; PS01209; LDLRA\_1; 8.  
DR PROSITE; PS00068; LDLRA\_2; 8.  
KW Glycoprotein; VLDL; Cholesterol metabolism; Lipid transport;  
KW Endocytosis; Coated pits; Transmembrane; Receptor; Signal;  
KW EGF-like domain; Repeat.  
FT SIGNAL 1 27  
FT CHAIN 28 873  
FT DOMAIN 28 797  
FT TRANSMEM 798 819  
FT DOMAIN 820 873  
FT DOMAIN 31 69  
FT DOMAIN 70 110  
FT DOMAIN 111 151  
FT DOMAIN 152 190  
FT DOMAIN 191 231  
FT DOMAIN 237 275  
FT DOMAIN 276 314  
FT DOMAIN 316 355  
FT DOMAIN 356 391  
FT DOMAIN 396 431  
FT REPEAT 439 480  
FT REPEAT 481 524  
FT REPEAT 525 567  
FT REPEAT 568 611  
FT REPEAT 612 654  
FT REPEAT 655 696  
FT DOMAIN 702 750  
FT DOMAIN 751 790  
FT SITE 832 837  
FT DISULFID 33 45  
FT DISULFID 40 58  
FT DISULFID 52 67  
FT DISULFID 72 84  
FT DISULFID 79 97  
FT DISULFID 91 108  
FT DISULFID 113 127  
FT DISULFID 120 140  
FT DISULFID 134 149  
FT DISULFID 154 166  
FT DISULFID 161 179  
FT DISULFID 173 188  
FT DISULFID 193 205  
FT DISULFID 200 218  
FT DISULFID 212 229  
FT DISULFID 239 251  
FT DISULFID 246 264  
FT DISULFID 258 273  
FT DISULFID 278 290  
FT DISULFID 285 303  
FT DISULFID 297 312  
FT DISULFID 318 331  
FT DISULFID 326 344  
FT DISULFID 338 355  
FT DISULFID 360 371  
FT DISULFID 367 380  
FT DISULFID 382 394  
FT DISULFID 400 410  
FT DISULFID 406 419  
FT DISULFID 421 434  
FT DISULFID 706 719  
FT DISULFID 715 734  
FT DISULFID 736 749  
FT CARBOHYD 151 151  
FT CARBOHYD 765 765

FT CARBOHYD 781 781 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CONFLICT 161  
FT CONFLICT 262 P -> G (IN REF. 1).  
FT CONFLICT 297 P -> L (IN REF. 3).  
SQ SEQUENCE 873 AA; 96372 MW; 08F09F93825195CB CRC64;  
Query Match 36.5%; Score 184.5; DB 1; Length 873;  
Best Local Similarity 42.0%; Pred. No. 1.3e-10;  
Matches 34; Conservative 15; Mismatches 27; Indels 5; Gaps 3;  
QY 1 KTCSPKQACRDQITCISKGRWCRDPCDGSDEAPEICQSKAQRCQPQHNK-LQTE 59  
Db 70 KTCAESDFVCKNG-QCVNRRWQCDGDPDCDGSDESPQC---HVRTCRINEISGARST 125  
QY 60 LCVPMRLNGVQDCMDSDE 80  
Db 126 QCIPIVSWRCDCGDCDNGEDE 146  
RESULT 12  
SORL RABIT  
ID SORL RABIT STANDARD; PRT; 2213 AA.  
AC Q95209;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Sortilin-related receptor precursor (Sorting protein-related receptor  
containing LDLR class A repeats) (SorLA) (SorLA-1) (Low-density  
lipoprotein receptor relative with 11 ligand-binding repeats) (LDLR  
relative with 11 ligand-binding repeats) (LR11).  
GN SORL1.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_Taxid=9986;  
RN [1]\_Taxid=9986;  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=96394640; PubMed=8798746;  
RA Yamazaki H., Bujo H., Kusunoki J., Seimiya K., Kanaki T., Morisaki N.,  
RA Schneider W.J., Saito Y.;  
RT "Elements of neural adhesion molecules and a yeast vacuolar protein  
sorting receptor are present in a novel mammalian low density  
lipoprotein receptor family member.";  
RL J. Biol. Chem. 271:24761-24768(1996).  
CC -!- FUNCTION: Likely to be a multifunctional endocytic receptor, that  
may be implicated in the uptake of lipoproteins and of proteases.  
CC Binds LDL, the major cholesterol-carrying lipoprotein of plasma,  
CC and transports it into cells by endocytosis. Binds the receptor-  
CC associated protein (RAP). Could play a role in cell-cell  
CC interaction (By similarity).  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).  
CC -!- TISSUE SPECIFICITY: Expressed abundantly in brain, in particular  
CC the hippocampus, dentate gyrus, and cerebral cortex, and is  
CC present at significant levels in liver, adrenal glands and  
CC testis.  
CC -!- PTM: THE PROPEPTIDE REMOVED IN THE N-TERMINUS MAY BE CLEAVED BY  
CC FURIN OR HOMOLOGOUS PROTEASES (BY SIMILARITY).  
CC -!- SIMILARITY: Contains 5 BNR repeats.  
CC -!- SIMILARITY: Contains 1 EGF-like domain.  
CC -!- SIMILARITY: Contains 11 LDL-receptor class A domains.  
CC -!- SIMILARITY: Contains 6 fibronectin type III domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D86350; BAAL3075.1; -.  
DR HSSP; P01130; 1AJJ.



RA Hjaelm G., Murray E., Crumley G., Harazim W., Lundgren S., Onyango I.,  
RA Ek B., Larsson M., Juhlin C., Hellman P., Davis H., Akerstrom G.,  
RA Rask L., Morse B.;  
RT "Cloning and sequencing of human gp330, a Ca(2+)-binding receptor  
with potential intracellular signaling properties.";  
RL Eur. J. Biochem. 239:132-137(1996).  
RN [2]  
RN SEQUENCE OF 2705-4453 FROM N.A.  
RC TISSUE=Kidney;  
RA Knaak C., Agraves W.S.;  
RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RN SEQUENCE OF 3933-4453 FROM N.A.  
RP TISSUE=Kidney;  
RX MEDLINE=95048397; PubMed=7953795;  
RA Korenberg J.R., Agraves K.M., Chen X.N., Tran H.,  
RA Strickland D.K., Agraves W.S.;  
RT "Chromosomal localization of human genes for the LDL receptor family  
member glycoprotein 330 (LRP2) and its associated protein RAP  
(LRPAP1).";  
RT LRPAP1.;  
RL Genomics 22:88-93(1994).  
RN [4]  
RN SEQUENCE OF 4139-4406 FROM N.A.  
RP MEDLINE=94244704; PubMed=8187828;  
RA Lundgren S., Hjaelm G., Hellman P., Ek B., Juhlin C., Rastad J.,  
RA Klarskog L., Akerstrom G., Rask L.;  
RT "A protein involved in calcium sensing of the human parathyroid and  
placental cytotrophoblast cells belongs to the LDL-receptor protein  
superfamily.";  
RL Exp. Cell Res. 212:344-350(1994).  
RN [5]  
RN FUNCTION  
RX MEDLINE=95286588; PubMed=7769901;  
RA Kounnas M.Z., Loukinova E.B., Stefansson S., Harmony J.A.K.,  
RA Brewer B.H., Strickland D.K., Agraves W.S.;  
RT "Identification of glycoprotein 330 as an endocytic receptor for  
apolipoprotein J/clusterin.";  
RL J. Biol. Chem. 270:13070-13075(1995).  
CC -1- FUNCTION: Binds specifically clusterin with high affinity, but  
also ligands in common with other family members: plasminogen,  
extracellular matrix components, plasminogen activator-plasminogen  
activator inhibitor type I complex, apolipoprotein E-enriched  
beta-VLDL, lipoprotein lipase, lactoferrin and calcium.  
CC -1- FUNCTION: Receptor-mediated uptake of polybasic drugs such as  
apronin, aminoglycosides and polymyxin B (By similarity).  
CC -1- FUNCTION: May participate in regulation of parathyroid-hormone and  
para-thyroid-hormone-related protein release.  
CC -1- SUBUNIT: Forms a multimeric complex together with a receptor-  
associated protein (RAP). Binds to ankyrin-repeat family A protein  
2 (ANKRA2) (By similarity).  
CC -1- TISSUE SPECIFICITY: Absorptive epithelia, including renal  
proximal tubules.  
CC -1- SIMILARITY: Contains 36 LDL-receptor class A domains.  
CC -1- SIMILARITY: Contains 37 LDL-receptor class B domains.  
CC -1- SIMILARITY: Contains 17 EGF-like domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U33837; BAB1649.1; --  
DR EMBL; U04441; AAB02882.1; --  
DR EMBL; S73145; AAB30825.1; --  
DR PIR; I53413; I53413.  
DR HSSP; Q07954; 1CR8.  
DR Genew; HGNC:6694; LRP2.  
DR MIM; 600073; --  
DR GO; GO:0005764; C:lysosome; TAS.

DR GO; GO:0006629; P:lipid metabolism; TAS.  
DR GO; GO:0006486; P:protein amino acid glycosylation; TAS.  
DR GO; GO:0006898; P:receptor mediated endocytosis; TAS.  
DR InterPro; IPR000152; Asx hydroxyl\_S.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR006209; EGF\_like.  
DR InterPro; IPR002172; LDL\_receptor\_A.  
DR InterPro; IPR000033; Ldl\_receptor\_rep.  
DR Pfam; PF00057; ldl\_recept\_a; 36.  
DR Pfam; PF00058; ldl\_recept\_b; 37.  
DR PRINTS; PR00261; LDLRECEPTOR.  
DR SMART; SM00179; EGF\_CA; 3.  
DR SMART; SM00192; LDLa; 36.  
DR SMART; SM00135; LY; 33.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 4.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS01186; EGF\_2; 9.  
DR PROSITE; PS00026; EGF\_3; 6.  
DR PROSITE; PS01187; EGF\_CA; 3.  
DR PROSITE; PS01209; LDLRA\_1; 31.  
DR PROSITE; PS00068; LDLRA\_2; 36.  
DR GlycoProtein; Repeat; Endocytosis; Coated pits; Transmembrane;  
KW Receptor; EGF-like domain; Signal; SH3-binding; Polymorphism.  
FT SIGNAL 1 25  
FT CHAIN 26 4555  
FT DOMAIN 26 4423  
FT TRANSMEM 4424 4446  
FT DOMAIN 4447 4555  
FT DOMAIN 26 64  
FT DOMAIN 65 105  
FT DOMAIN 106 144  
FT DOMAIN 145 181  
FT DOMAIN 182 219  
FT DOMAIN 220 258  
FT DOMAIN 264 308  
FT DOMAIN 309 347  
FT DOMAIN 348 386  
FT DOMAIN 436 477  
FT DOMAIN 479 520  
FT DOMAIN 522 567  
FT DOMAIN 569 612  
FT DOMAIN 613 653  
FT DOMAIN 659 705  
FT DOMAIN 753 794  
FT DOMAIN 796 836  
FT DOMAIN 838 880  
FT DOMAIN 882 924  
FT DOMAIN 970 1014  
FT DOMAIN 1024 1062  
FT DOMAIN 1065 1103  
FT DOMAIN 1107 1145  
FT DOMAIN 1147 1185  
FT DOMAIN 1186 1224  
FT DOMAIN 1228 1268  
FT DOMAIN 1269 1307  
FT DOMAIN 1310 1350  
FT DOMAIN 1349 1389  
FT DOMAIN 1390 1429  
FT DOMAIN 1478 1519  
FT DOMAIN 1521 1562  
FT DOMAIN 1566 1608  
FT DOMAIN 1610 1653  
FT DOMAIN 1655 1695  
FT DOMAIN 1700 1741  
FT DOMAIN 1790 1831  
FT DOMAIN 1833 1881  
FT DOMAIN 1893 1929  
FT DOMAIN 1931 1971  
FT DOMAIN 1972 2012  
FT DOMAIN 2018 2059  
FT DOMAIN 2107 2155

EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).

```

FT DOMAIN 2157 2200 LDL-RECEPTOR CLASS B 21.
FT DOMAIN 2202 2244 LDL-RECEPTOR CLASS B 22.
FT DOMAIN 2246 2288 LDL-RECEPTOR CLASS B 23.
FT DOMAIN 2290 2331 LDL-RECEPTOR CLASS B 24.
FT DOMAIN 2342 2383 EGF-LIKE 9.
FT DOMAIN 2431 2476 LDL-RECEPTOR CLASS B 25.
FT DOMAIN 2478 2517 LDL-RECEPTOR CLASS B 26.
FT DOMAIN 2519 2561 LDL-RECEPTOR CLASS B 27.
FT DOMAIN 2563 2603 LDL-RECEPTOR CLASS B 28.
FT DOMAIN 2604 2645 LDL-RECEPTOR CLASS B 29.
FT DOMAIN 2651 2693 EGF-LIKE 10.
FT DOMAIN 2698 2737 LDL-RECEPTOR CLASS A 16.
FT DOMAIN 2739 2777 LDL-RECEPTOR CLASS A 17.
FT DOMAIN 2778 2819 LDL-RECEPTOR CLASS A 18.
FT DOMAIN 2820 2861 LDL-RECEPTOR CLASS A 19.
FT DOMAIN 2862 2901 LDL-RECEPTOR CLASS A 20.
FT DOMAIN 2904 2945 LDL-RECEPTOR CLASS A 21.
FT DOMAIN 2946 2990 LDL-RECEPTOR CLASS A 22.
FT DOMAIN 2991 3029 LDL-RECEPTOR CLASS A 23.
FT DOMAIN 3030 3070 LDL-RECEPTOR CLASS A 24.
FT DOMAIN 3073 3110 LDL-RECEPTOR CLASS A 25.
FT DOMAIN 3111 3151 EGF-LIKE 11.
FT DOMAIN 3152 3192 EGF-LIKE 12.
FT DOMAIN 3193 3239 LDL-RECEPTOR CLASS B 30.
FT DOMAIN 3240 3282 LDL-RECEPTOR CLASS B 31.
FT DOMAIN 3283 3333 LDL-RECEPTOR CLASS B 32.
FT DOMAIN 3334 3375 LDL-RECEPTOR CLASS B 33.
FT DOMAIN 3376 3419 LDL-RECEPTOR CLASS B 34.
FT DOMAIN 3420 3465 EGF-LIKE 13.
FT DOMAIN 3466 3509 LDL-RECEPTOR CLASS A 26.
FT DOMAIN 3510 3550 LDL-RECEPTOR CLASS A 27.
FT DOMAIN 3551 3591 LDL-RECEPTOR CLASS A 28.
FT DOMAIN 3592 3632 LDL-RECEPTOR CLASS A 29.
FT DOMAIN 3633 3673 LDL-RECEPTOR CLASS A 30.
FT DOMAIN 3674 3716 LDL-RECEPTOR CLASS A 31.
FT DOMAIN 3717 3756 LDL-RECEPTOR CLASS A 32.
FT DOMAIN 3757 3795 LDL-RECEPTOR CLASS A 33.
FT DOMAIN 3796 3834 LDL-RECEPTOR CLASS A 34.
FT DOMAIN 3835 3880 LDL-RECEPTOR CLASS A 34.

Query Match 36.2%; Score 183; DB 1; Length 4655;
Best local similarity 42.9%; Pred. No. 8.2e-10;
Matches 36; Conservative 13; Mismatches 31; Indels 4; Gaps 3;

QY 1 KTCSPKQFACRQITICISKWRGDCRDCPDGSDPEAPICPSKACRQCPQNEHNCGLTEL 60
Db 1269 KTCPSVFHC-DNGNCIHRWLCDRNDGCDMSDE--KDCP-TQPRCPSCWQCLGHNI 1324

QY 61 CVPMSRLCNVQDCMDGSDGPHC 84
Db 1325 CYNLVVCDGIFDCPNGTDSPLC 1348

RESULT 14
LDLR_RABIT STANDARD; PRT; 837 AA.
AC P20063;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Low-density lipoprotein receptor (LDL receptor) (Fragment).
GN LDLR.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
CX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86208167; PubMed=3010466;
RA Yamamoto T., Bishop R.W., Brown M.S., Goldstein J.L., Russell D.W.;
RT "Deletion in cysteine-rich region of LDL receptor impedes transport
RT to cell surface in WHHL rabbit."
RL Science 232:1230-1237(1986).
CC -!- FUNCTION: Binds LDL, the major cholesterol-carrying lipoprotein of

```

```

CC plasma, and transports it into cells by endocytosis. In order to
CC be internalized, the receptor-ligand complexes must first cluster
CC into clathrin-coated pits.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: Contains 7 LDL-receptor class A domains.
CC -!- SIMILARITY: Contains 6 LDL-receptor class B domains.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M11501; AAA31383.1; -.
CC PIR; A29512; A29512.
CC HSP; P01130; LDL.
CC InterPro; IPR000152; Asx_hydroxyl_S.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR006209; EGF_Like.
CC InterPro; IPR002172; LDL_receptor_A.
CC InterPro; IPR000033; Ldl_receptor_rep.
CC Pfam; PF00008; EGF_2.
CC Pfam; PF00057; ldl_recept_a; 7.
CC Pfam; PF00058; ldl_recept_b; 5.
CC PRINTS; PR00261; LDLRECEPTOR.
CC SMART; SM00179; EGF_CA; 1.
CC SMART; SM00192; LDLa; 7.
CC SMART; SM00135; LY; 5.
CC PROSITE; PS00010; ASX_HYDROXYL; 1.
CC PROSITE; PS00022; EGF_1; FALSE_NEG.
CC PROSITE; PS01186; EGF_2; 2.
CC PROSITE; PS00026; EGF_3; 2.
CC PROSITE; PS01187; EGF_CA; 1.
CC PROSITE; PS01209; LDLRA_1; 7.
CC PROSITE; PS00068; LDLRA_2; 7.
CC Glycoprotein; LDL; Cholesterol metabolism; Lipid transport;
CC Endocytosis; Coated pits; Transmembrane; Receptor; EGF-like domain;
CC Repeat.
FT NON TER 1 765 EXTRACELLULAR (POTENTIAL).
FT DOMAIN <1 766 787 POTENTIAL.
FT TRANSMEM 788 837 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 12 52 LDL-RECEPTOR CLASS A 1.
FT DOMAIN 53 93 LDL-RECEPTOR CLASS A 2.
FT DOMAIN 94 132 LDL-RECEPTOR CLASS A 3.
FT DOMAIN 133 173 LDL-RECEPTOR CLASS A 4.
FT DOMAIN 182 220 LDL-RECEPTOR CLASS A 5.
FT DOMAIN 221 259 LDL-RECEPTOR CLASS A 6.
FT DOMAIN 261 300 LDL-RECEPTOR CLASS A 7.
FT DOMAIN 301 340 EGF-LIKE 1.
FT DOMAIN 341 380 EGF-LIKE 2. CALCIUM-BINDING (POTENTIAL).
FT REPEAT 384 425 LDL-RECEPTOR CLASS B 1.
FT REPEAT 426 471 LDL-RECEPTOR CLASS B 2.
FT REPEAT 473 515 LDL-RECEPTOR CLASS B 3.
FT REPEAT 516 559 LDL-RECEPTOR CLASS B 4.
FT REPEAT 560 602 LDL-RECEPTOR CLASS B 5.
FT REPEAT 603 644 LDL-RECEPTOR CLASS B 6.
FT DOMAIN 650 699 EGF-LIKE 3.
FT SITE 800 805 ENDOCYTOSIS SIGNAL (POTENTIAL).
FT DISULFID 14 26 BY SIMILARITY.
FT DISULFID 21 39 BY SIMILARITY.
FT DISULFID 33 59 BY SIMILARITY.
FT DISULFID 55 69 BY SIMILARITY.
FT DISULFID 62 82 BY SIMILARITY.
FT DISULFID 76 91 BY SIMILARITY.
FT DISULFID 96 108 BY SIMILARITY.
FT DISULFID 103 121 BY SIMILARITY.
FT DISULFID 115 130 BY SIMILARITY.
FT DISULFID 135 147 BY SIMILARITY.
FT DISULFID 142 160 BY SIMILARITY.

```



	FT	DISULFID	152	167	BY SIMILARITY.
	FT	DISULFID	172	184	BY SIMILARITY.
	FT	DISULFID	179	197	BY SIMILARITY.
	FT	DISULFID	191	206	BY SIMILARITY.
	FT	DISULFID	211	223	BY SIMILARITY.
	FT	DISULFID	218	236	BY SIMILARITY.
	FT	DISULFID	230	247	BY SIMILARITY.
	FT	DISULFID	237	269	BY SIMILARITY.
	FT	DISULFID	264	282	BY SIMILARITY.
	FT	DISULFID	276	291	BY SIMILARITY.
	FT	DISULFID	296	308	BY SIMILARITY.
	FT	DISULFID	303	321	BY SIMILARITY.
	FT	DISULFID	315	330	BY SIMILARITY.
	FT	DISULFID	336	349	BY SIMILARITY.
	FT	DISULFID	344	362	BY SIMILARITY.
	FT	DISULFID	356	373	BY SIMILARITY.
	FT	DISULFID	378	389	BY SIMILARITY.
	FT	DISULFID	385	398	BY SIMILARITY.
	FT	DISULFID	400	412	BY SIMILARITY.
	FT	DISULFID	418	428	BY SIMILARITY.
	FT	DISULFID	424	437	BY SIMILARITY.
	FT	DISULFID	439	452	BY SIMILARITY.
	FT	DISULFID	726	739	BY SIMILARITY.
	FT	DISULFID	735	754	BY SIMILARITY.
	FT	DISULFID	756	769	BY SIMILARITY.
	FT	CARBONYD	169	169	N-LINKED (GLCNAC. . .) (POTENTIAL).
	FT	CARBONYD	773	773	N-LINKED (GLCNAC. . .) (POTENTIAL).
	SQ	SEQUENCE	863 AA;	94904 MW;	0672A8748F9A2245 CRC64;
		Query Match	36.1%;	Score 182.5;	DB 1; Length 863;
		Best Local Similarity	46.2%;	Fred. No. 2e-10;	
		Matches 36;	Conservative 11;	Mismatches 24;	Indels 7; Gaps 4;
	QY	3	CSEPKGFACRDQITCISKGWRCDGERDCPDGSDEAPICPSKAQRCPNHNCLGTETCV 62		
					:::::
	Dd	257	CSTSEVQGSG-ECHIKKRWRCGDPCDKDSGIN--CP---SPTCRDPDFCEDGN-CI 309		
					:::::
	QY	63	PMSRLCNVGQDMGSDS 80		
					:::::
	Dd	310	HGSROCGVRDCLDGTDE 327		
					:::::

Search completed: September 17, 2004, 11:15:51  
Job time : 26 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2004, 11:09:54 ; Search time 116 Seconds  
(without alignments)  
233.919 Million cell updates/sec

Title: US-09-625-137-21

Perfect score: 506

Sequence: 1 KTCSPKQFACRDQITCISK.....LCNGVDCMGDSDEGPHCRE 86

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	506	100.0	292	4 Q86SW0	Q86sw0 homo sapien
2	506	100.0	296	4 Q727K9	Q727k9 homo sapien
3	486	96.0	4545	11 Q912X7	Q912x7 mus musculu
4	486	96.0	4545	11 Q920Y4	Q920y4 mus musculu
5	486	96.0	4545	11 Q61291	Q61291 mus musculu
6	280	55.3	881	4 Q8WY31	Q8wy31 homo sapien
7	280	55.3	4599	4 Q9NZR2	Q9n2r2 homo sapien
8	270	53.4	4599	11 Q9J118	Q9j118 mus musculu
9	209	41.3	883	5 Q46131	Q46131 locusta nig
10	207.5	41.0	883	5 Q9VBN1	Q9vbn1 drosophila
11	205.5	40.6	4699	5 Q9V383	Q9v383 drosophila
12	204.5	40.4	891	5 Q7YW57	Q7yw57 aedes aegypt
13	204.5	40.4	1156	5 Q963T3	Q963t3 aedes aegypt
14	203	40.1	1537	4 Q8WY29	Q8wy29 homo sapien
15	202	39.9	739	5 Q8IGR9	Q8igr9 drosophila
16	202	39.9	826	5 Q86B77	Q86b77 drosophila

#### ALIGNMENTS

##### RESULT 1

Q86SW0 PRELIMINARY; PRT; 292 AA.

AC Q86SW0;

DT 01-JUN-2003 (TRENBLrel. 24, Created)

DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)

DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)

DE Similar to low density lipoprotein-related protein 1 (Alpha-2-macroglobulin receptor).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Melanoma;

RA Strausberg R.;

EL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC045107; AAH45107.1; -.

DR GO; GO:0005509; F:calcium ion binding; IEA.

DR InterPro; IPR000152; Asx\_hydroxyl\_S.

DR InterPro; IPR001881; EGF\_Ca.

DR InterPro; IPR006209; EGF-like.

DR InterPro; IPR006210; EGF.

DR InterPro; IPR002172; LDL\_receptor\_A.

DR Pfam; PF00008; EGF; 2.

DR Pfam; PF00057; ldl\_recept\_a; 2.

DR PRINTS; PR00261; LDLRECEPTOR.

DR SMART; SM00181; EGF; 2.

DR SMART; SM00179; EGF\_CA; 2.

DR SMART; SM00192; LDLa; 2.

DR PROSITE; PS00010; ASX\_HYDROXYL; 1.

DR PROSITE; PS01186; EGF\_2; 1.

DR PROSITE; PS01187; EGF\_CA; 1.

DR PROSITE; PS01209; LDLRA\_1; 2.

DR PROSITE; PS50068; LDLRA\_2; 2.

DR Lipoproteins.

SQ SEQUENCE 292 AA; 31631 MW; 82C39315F1BCE24 CRC64;

Q7YTz6 drosophila  
Q7YU01 drosophila  
Q9VBN2 drosophila  
Q95QH2 caenorhabdi  
Q9VY56 drosophila  
Q86P52 drosophila  
Q802V2 brachydanio  
Q96NT6 homo sapien  
Q921B6 mus musculu  
Q86V27 homo sapien  
Q141L4 homo sapien  
Q90W12 oncorhynchu  
Q7C2X3 gallus gall  
Q7ZTG7 oreochromis  
Q7SXV0 brachydanio  
Q99876 homo sapien  
Q14968 homo sapien  
O01552 caenorhabdi  
Q91Z11 mus musculu  
Q9VCT0 mus musculu  
O01768 caenorhabdi  
Q9VTE6 rattus sp.  
Q8CPS5 mus musculu  
Q924X6 mus musculu  
Q9V156 mus musculu  
O02660 bos taurus  
Q8SPM4 bos taurus  
O42126 xenopus lae  
Q8T4N8 penaeus sem

```

Query Match      100.0%; Score 506; DB 4; Length 292;
Best Local Similarity 100.0%; Pred. No. 5.5e-55;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRQDITCISKWRCDCGDCPDGSDPAEIPCQSKAQCQPNHNCGLGTEL 60
DB 25 KTCSPKQFACRQDITCISKWRCDCGDCPDGSDPAEIPCQSKAQCQPNHNCGLGTEL 84

QY 61 CVPMSRLCNGVQDCMDSDEGPHCRE 86
DB 85 CVPMSRLCNGVQDCMDSDEGPHCRE 110

RESULT 2
Q727K9 PRELIMINARY; PRT; 296 AA.
AC Q727K9;
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Reingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McWay P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.O., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
PL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Strausberg R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC052593; AAH52593.1; -.
KW Hypothetical protein.
SQ SEQUENCE 296 AA; 32226 MW; C3FEF9F2F13E5854 CRC64;

Query Match      100.0%; Score 506; DB 4; Length 296;
Best Local Similarity 100.0%; Pred. No. 5.6e-55;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRQDITCISKWRCDCGDCPDGSDPAEIPCQSKAQCQPNHNCGLGTEL 60
DB 25 KTCSPKQFACRQDITCISKWRCDCGDCPDGSDPAEIPCQSKAQCQPNHNCGLGTEL 84

QY 61 CVPMSRLCNGVQDCMDSDEGPHCRE 86
DB 85 CVPMSRLCNGVQDCMDSDEGPHCRE 110

RESULT 3
Q91ZX7 PRELIMINARY; PRT; 4545 AA.
AC Q91ZX7;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Lipoprotein receptor-related protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBA;
RA Roebroek A.J.M.;
RT "LFP cDNA sequence of CBA mouse strain.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF367720; AAL09566.1; -.
DR MGD; MGI_96828; Lxpl.
DR GO; GO:0016020; C-membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR000152; ASX_hydroxyl_S.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000033; LDL_receptor_rep.
DR Pfam; PF00008; EGF; 14.
DR Pfam; PF00057; ldl_recept_a; 31.
DR Pfam; PF00058; ldl_recept_b; 33.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00179; EGF_CA; 3.
DR SMART; SM00192; LDLA; 31.
DR SMART; SM00135; LY; 35.
DR PROSITE; PS00010; ASX_HYDROXYL; 3.
DR PROSITE; PS00022; EGF_1; 5.
DR PROSITE; PS01186; EGF_2; 8.
DR PROSITE; PS01187; EGF_CA; 2.
DR PROSITE; PS01209; LDLRA_1; 27.
DR PROSITE; PS00068; LDLRA_2; 29.
KW EGF-like domain; Receptor.
SQ SEQUENCE 4545 AA; 504730 MW; 9904CF5DF5EB333E CRC64;

Query Match      96.0%; Score 486; DB 11; Length 4545;
Best Local Similarity 95.3%; Pred. No. 2.9e-51;
Matches 82; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTCSPKQFACRQDITCISKWRCDCGDCPDGSDPAEIPCQSKAQCQPNHNCGLGTEL 60
DB 26 KTCSPKQFACRQDITCISKWRCDCGDCPDGSDPAEIPCQSKAQCQPNHNCGLGTEL 85

QY 61 CVPMSRLCNGVQDCMDSDEGPHCRE 86
DB 86 CVPMSRLCNGVQDCMDSDEGPHCRE 111

RESULT 4
Q920Y4 PRELIMINARY; PRT; 4545 AA.
AC Q920Y4;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Lipoprotein receptor-related protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129P3/J;
RA Roebroek A.J.M.;

```



"Lrp1 exon sequences of 129P3/J mouse strain."; submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.

```

RT      DR EMBL; AF369477; AAL09567.1; JOINED.
BL      DR EMBL; AF369477; AAL09567.1; JOINED.
DR      DR EMBL; AF369389; AAL09567.1; JOINED.
DR      DR EMBL; AF369390; AAL09567.1; JOINED.
DR      DR EMBL; AF369391; AAL09567.1; JOINED.
DR      DR EMBL; AF369392; AAL09567.1; JOINED.
DR      DR EMBL; AF369393; AAL09567.1; JOINED.
DR      DR EMBL; AF369394; AAL09567.1; JOINED.
DR      DR EMBL; AF369395; AAL09567.1; JOINED.
DR      DR EMBL; AF369396; AAL09567.1; JOINED.
DR      DR EMBL; AF369397; AAL09567.1; JOINED.
DR      DR EMBL; AF369398; AAL09567.1; JOINED.
DR      DR EMBL; AF369399; AAL09567.1; JOINED.
DR      DR EMBL; AF369400; AAL09567.1; JOINED.
DR      DR EMBL; AF369401; AAL09567.1; JOINED.
DR      DR EMBL; AF369402; AAL09567.1; JOINED.
DR      DR EMBL; AF369403; AAL09567.1; JOINED.
DR      DR EMBL; AF369404; AAL09567.1; JOINED.
DR      DR EMBL; AF369405; AAL09567.1; JOINED.
DR      DR EMBL; AF369406; AAL09567.1; JOINED.
DR      DR EMBL; AF369407; AAL09567.1; JOINED.
DR      DR EMBL; AF369408; AAL09567.1; JOINED.
DR      DR EMBL; AF369409; AAL09567.1; JOINED.
DR      DR EMBL; AF369410; AAL09567.1; JOINED.
DR      DR EMBL; AF369411; AAL09567.1; JOINED.
DR      DR EMBL; AF369412; AAL09567.1; JOINED.
DR      DR EMBL; AF369413; AAL09567.1; JOINED.
DR      DR EMBL; AF369414; AAL09567.1; JOINED.
DR      DR EMBL; AF369415; AAL09567.1; JOINED.
DR      DR EMBL; AF369416; AAL09567.1; JOINED.
DR      DR EMBL; AF369417; AAL09567.1; JOINED.
DR      DR EMBL; AF369418; AAL09567.1; JOINED.
DR      DR EMBL; AF369419; AAL09567.1; JOINED.
DR      DR EMBL; AF369420; AAL09567.1; JOINED.
DR      DR EMBL; AF369421; AAL09567.1; JOINED.
DR      DR EMBL; AF369422; AAL09567.1; JOINED.
DR      DR EMBL; AF369423; AAL09567.1; JOINED.
DR      DR EMBL; AF369424; AAL09567.1; JOINED.
DR      DR EMBL; AF369425; AAL09567.1; JOINED.
DR      DR EMBL; AF369426; AAL09567.1; JOINED.
DR      DR EMBL; AF369427; AAL09567.1; JOINED.
DR      DR EMBL; AF369428; AAL09567.1; JOINED.
DR      DR EMBL; AF369429; AAL09567.1; JOINED.
DR      DR EMBL; AF369430; AAL09567.1; JOINED.
DR      DR EMBL; AF369431; AAL09567.1; JOINED.
DR      DR EMBL; AF369432; AAL09567.1; JOINED.
DR      DR EMBL; AF369433; AAL09567.1; JOINED.
DR      DR EMBL; AF369434; AAL09567.1; JOINED.
DR      DR EMBL; AF369435; AAL09567.1; JOINED.
DR      DR EMBL; AF369436; AAL09567.1; JOINED.
DR      DR EMBL; AF369437; AAL09567.1; JOINED.
DR      DR EMBL; AF369438; AAL09567.1; JOINED.
DR      DR EMBL; AF369439; AAL09567.1; JOINED.
DR      DR EMBL; AF369440; AAL09567.1; JOINED.
DR      DR EMBL; AF369441; AAL09567.1; JOINED.
DR      DR EMBL; AF369442; AAL09567.1; JOINED.
DR      DR EMBL; AF369443; AAL09567.1; JOINED.
DR      DR EMBL; AF369444; AAL09567.1; JOINED.
DR      DR EMBL; AF369445; AAL09567.1; JOINED.
DR      DR EMBL; AF369446; AAL09567.1; JOINED.
DR      DR EMBL; AF369447; AAL09567.1; JOINED.
DR      DR EMBL; AF369448; AAL09567.1; JOINED.
DR      DR EMBL; AF369449; AAL09567.1; JOINED.
DR      DR EMBL; AF369450; AAL09567.1; JOINED.
DR      DR EMBL; AF369451; AAL09567.1; JOINED.
DR      DR EMBL; AF369452; AAL09567.1; JOINED.
DR      DR EMBL; AF369453; AAL09567.1; JOINED.
DR      DR EMBL; AF369454; AAL09567.1; JOINED.
DR      DR EMBL; AF369455; AAL09567.1; JOINED.
DR      DR EMBL; AF369456; AAL09567.1; JOINED.
DR      DR EMBL; AF369457; AAL09567.1; JOINED.
DR      DR EMBL; AF369458; AAL09567.1; JOINED.

DR      DR EMBL; AF369459; AAL09567.1; JOINED.
DR      DR EMBL; AF369460; AAL09567.1; JOINED.
DR      DR EMBL; AF369461; AAL09567.1; JOINED.
DR      DR EMBL; AF369462; AAL09567.1; JOINED.
DR      DR EMBL; AF369463; AAL09567.1; JOINED.
DR      DR EMBL; AF369464; AAL09567.1; JOINED.
DR      DR EMBL; AF369465; AAL09567.1; JOINED.
DR      DR EMBL; AF369466; AAL09567.1; JOINED.
DR      DR EMBL; AF369467; AAL09567.1; JOINED.
DR      DR EMBL; AF369468; AAL09567.1; JOINED.
DR      DR EMBL; AF369469; AAL09567.1; JOINED.
DR      DR EMBL; AF369470; AAL09567.1; JOINED.
DR      DR EMBL; AF369471; AAL09567.1; JOINED.
DR      DR EMBL; AF369472; AAL09567.1; JOINED.
DR      DR EMBL; AF369473; AAL09567.1; JOINED.
DR      DR EMBL; AF369474; AAL09567.1; JOINED.
DR      DR EMBL; AF369475; AAL09567.1; JOINED.
DR      DR EMBL; AF369476; AAL09567.1; JOINED.
DR      DR GO; GO:0016020; C:membrane; IEA.
DR      DR GO; GO:0005509; P:calcium ion binding; IEA.
DR      DR GO; GO:0004872; P:receptor activity; IEA.
DR      DR InterPro; IPR000152; ASX_HYDROXYL_S.
DR      DR InterPro; IPR001881; EGF_CA.
DR      DR InterPro; IPR006209; EGF_LIKE.
DR      DR InterPro; IPR002172; LDL_receptor_A.
DR      DR InterPro; IPR000033; LDL_receptor_rep.
DR      DR Pfam; PF00008; EGF_14.
DR      DR Pfam; PF00057; LDL_recept_a; 31.
DR      DR Pfam; PF00058; LDL_recept_b; 33.
DR      DR PRINTS; PR00261; LDLRECEPTOR.
DR      DR SMART; SM00179; EGF_CA; 3.
DR      DR SMART; SM00192; LDLa; 31.
DR      DR SMART; SM00135; LY; 35.
DR      DR PROSITE; PS00010; ASX_HYDROXYL; 3.
DR      DR PROSITE; PS00022; EGF_1; 5.
DR      DR PROSITE; PS01186; EGF_2; 8.
DR      DR PROSITE; PS01187; EGF_CA; 2.
DR      DR PROSITE; PS01209; LDLRA_1; 27.
DR      DR PROSITE; PS00068; LDLRA_2; 29.
KW      EGF-like domain; Receptor.
SQ      SEQUENCE 4545 AA; 504760 MW; C92B69611F7A5212 CRC64;

Query Match      96.0%; Score 486; DB 11; Length 4545;
Best Local Similarity 95.3%; Pred. No. 2.9e-51;
Matches 82; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 KTCSPKQFACRDQITCISKGWRCDGRCDCPDGSDPEAFCPSKAQRCPNHNCLGTGL 60
        |||||
Db      26 KTCSPKQFACRDQITCISKGWRCDGRCDCPDGSDPEAFCPSKAQRCPNHNCLGTGL 85
        |||||

QY      61 CVPMSRLCNGVQDCMDGSDGPHCRE 86
        |||||
Db      86 CVPMSRLCNGIQDCMDGSDGHAHCRE 111
        |||||

RESULT 5
Q61291 ID Q61291 PRELIMINARY; PRT; 4545 AA.
AC Q61291;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE AM2 receptor.
GN LRP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]__TaxID=10090;
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93250049; PubMed=8485155;
RA van Leuven F.;
```

"Molecular cloning and sequencing of the Arawine L2M-receptor cDNA.";  
 RL Biochim. Biophys. Acta 1173:71-74(1993).  
 DR EMBL; X67469; CAA47817.1; -.  
 DR PIR; S25111; S25111.  
 DR HSP; O07954; ICR8.  
 DR MGD; MGI:96828; Lrp1.  
 DR GO; GO:0016020; C-membrane; IEA.  
 DR GO; GO:0005509; F:calcium ion binding; IEA.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR InterPro; IPR000152; Asx\_hydroxyl\_S.  
 DR InterPro; IPR001881; EGF\_Ca.  
 DR InterPro; IPR006209; EGF\_like.  
 DR InterPro; IPR002172; LDL\_receptor\_A.  
 DR InterPro; IPR000033; LDL\_receptor\_rep.  
 DR Pfam; PF00008; EGF; 15.  
 DR Pfam; PF00057; ldl\_recept\_a; 31.  
 DR Pfam; PF00058; ldl\_recept\_b; 34.  
 DR PRINTS; PR00261; LDLRECEPTOR.  
 DR SMART; SM00179; EGF CA; 3.  
 DR SMART; SM00192; LDLa; 31.  
 DR SMART; SM00135; LY; 33.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 3.  
 DR PROSITE; PS00022; EGF\_1; 5.  
 DR PROSITE; PS01186; EGF\_2; 8.  
 DR PROSITE; PS01187; EGF\_CA; 2.  
 DR PROSITE; PS01209; LDLRA\_1; 26.  
 DR PROSITE; PS00088; LDLRA\_2; 29.  
 DR EGF-like domain; Receptor.  
 KW EGF-like domain; Receptor.  
 SQ SEQUENCE 4545 AA; 504754 MW; A9263A81CF45E56D CRC64;  
 Query Match 96.0%; Score 486; DB 11; Length 4545;  
 Best Local Similarity 95.3%; Pred. No. 2.9e-51;  
 Matches 82; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 KTSPKQFACRDQITCISKWRCDGERDCPDGSDPEAIPCQSKAQRCPNHNCLGTEL 60  
 DB 26 KTSPKQFACRDQITCISKWRCDGERDCPDGSDPEAIPCQSKAQRCPNHNCLGTEL 85  
 QY 61 CVPMRLCNGVQDCMGSDGEGPHCRE 86  
 DB 86 CVPMRLCNGVQDCMGSDGEGPHCRE 111  
 RESULT 6  
 ID Q8WY31 PRELIMINARY; PRT; 881 AA.  
 AC Q8WY31;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Candidate tumor suppressor protein (fragment).  
 GN LRP1B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20487554; PubMed=11031110;  
 RA Liu C.-X., Musco S., Lisitsyna N.M., Yaklichkin S.Y., Lisitsyn N.A.;  
 RT "Genomic organization of a new candidate tumor suppressor gene,  
 LRP1B";  
 RL Genomics 69:271-274(2000).  
 DR EMBL; AF283334; AAL38107.1; -.  
 DR EMBL; AF283327; AAL38107.1; JOINED.  
 DR EMBL; AF283328; AAL38107.1; JOINED.  
 DR EMBL; AF283329; AAL38107.1; JOINED.  
 DR EMBL; AF283330; AAL38107.1; JOINED.  
 DR EMBL; AF283331; AAL38107.1; JOINED.  
 DR EMBL; AF283332; AAL38107.1; JOINED.  
 DR EMBL; AF283333; AAL38107.1; JOINED.  
 DR EMBL; AF283334; AAL38107.1; JOINED.  
 DR EMBL; AF283335; AAL38107.1; JOINED.

DR EMBL; AF283336; AAL38107.1; JOINED.  
 DR EMBL; AF283337; AAL38107.1; JOINED.  
 DR EMBL; AF283338; AAL38107.1; JOINED.  
 DR EMBL; AF283339; AAL38107.1; JOINED.  
 DR EMBL; AF283340; AAL38107.1; JOINED.  
 DR EMBL; AF283341; AAL38107.1; JOINED.  
 DR GO; GO:0016020; C-membrane; IEA.  
 DR GO; GO:0005634; C-nucleus; IEA.  
 DR GO; GO:0005509; F:calcium ion binding; IEA.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR InterPro; IPR000152; Asx\_hydroxyl\_S.  
 DR InterPro; IPR001881; EGF\_Ca.  
 DR InterPro; IPR006209; EGF\_like.  
 DR InterPro; IPR002172; LDL\_receptor\_A.  
 DR InterPro; IPR000033; LDL\_receptor\_rep.  
 DR Pfam; PF00008; EGF; 3.  
 DR Pfam; PF00057; ldl\_recept\_a; 3.  
 DR Pfam; PF00058; ldl\_recept\_b; 7.  
 DR PRINTS; PR00261; LDLRECEPTOR.  
 DR SMART; SM00179; EGF CA; 1.  
 DR SMART; SM00192; LDLa; 3.  
 DR SMART; SM00135; LY; 9.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 1.  
 DR PROSITE; PS01186; EGF\_2; 2.  
 DR PROSITE; PS01187; EGF\_CA; 1.  
 DR PROSITE; PS01209; LDLRA\_1; 2.  
 DR PROSITE; PS00068; LDLRA\_2; 3.  
 DR PROSITE; PS00037; MYB\_1; 1.  
 DR EGF-like domain.  
 KW EGF-like domain.  
 FT NON TER 881  
 SQ SEQUENCE 881 AA; 99936 MW; 8325F9D984CADESA CRC64;  
 Query Match 55.3%; Score 280; DB 4; Length 881;  
 Best Local Similarity 53.6%; Pred. No. 2.7e-26;  
 Matches 45; Conservative 12; Mismatches 27; Indels 0; Gaps 0;  
 QY 3 CSPKQFACRDQITCISKWRCDGERDCPDGSDPEAIPCQSKAQRCPNHNCLGTEL 62  
 DB 32 CDPGEFLCHDHVTCVSWLSDGDPDSDSDSLDTCPEVEIKPLNHIACLGTKCV 91  
 QY 63 PMSRLCNGVQDCMGSDGEGPHCRE 86  
 DB 92 HSLQCLNGVLDCPDGDEGVHCQE 115  
 RESULT 7  
 ID Q9NZR2 PRELIMINARY; PRT; 4599 AA.  
 AC Q9NZR2;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Low density lipoprotein receptor related protein-deleted in  
 DE tumor.  
 GN LRPDIT.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20227253; PubMed=10766186;  
 RA Liu C.-X., Musco S., Lisitsyna N.M., Forgacs E., Minna J.D.,  
 RA Lisitsyn N.A.;  
 RT "LRP-DIT, a putative endocytic receptor gene, is frequently  
 RT inactivated in non-small cell lung cancer cell lines";  
 RL Cancer Res. 60:1961-1967(2000).  
 DR EMBL; AF176832; AAF70379.1; -.  
 DR HSP; O07954; ICR8.  
 DR Genew; HGNC:6693; LRP1B.  
 DR GO; GO:0015029; F:internalization receptor activity; TAS.  
 DR GO; GO:0015031; P:protein transport; TAS.

GO; GO:000698; P:receptor mediated endocytosis; TAS.  
DR InterPro; IPR000152; Asx hydroxyl\_S.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR006209; EGF\_like.  
DR InterPro; IPR002172; LDL\_receptor\_A.  
DR InterPro; IPR000033; LDL\_receptor\_rep.  
DR InterPro; IPR001005; Myb DNA binding.  
DR InterPro; IPR000169; SHprot\_acsite.  
DR Pfam; PF00008; EGF; 15.  
DR Pfam; PF00057; ldl\_recept\_a; 32.  
DR Pfam; PF00058; ldl\_recept\_b; 34.  
DR PRINTS; PR00261; LDLRECEPTOR.  
DR SMART; SM00179; EGF\_CA; 4.  
DR SMART; SM00192; LDLA; 32.  
DR PROSITE; PS00010; ASX HYDROXYL; 4.  
DR PROSITE; PS00022; EGF\_1; 5.  
DR PROSITE; PS01186; EGF\_2; 9.  
DR PROSITE; PS01187; EGF\_CA; 3.  
DR PROSITE; PS01209; LDLRA\_1; 27.  
DR PROSITE; PS00068; LDLRA\_2; 32.  
DR PROSITE; PS00037; MYB\_1; 1.  
DR PROSITE; PS00139; THIOLE PROTEASE\_CYS; 1.  
KW EGF-like domain; Lipoprotein; Receptor.  
SQ SEQUENCE 4599 AA; 515498 MW; 9A97A555F1FAAE66 CRC64;  
  
Query Match 55.3%; Score 280; DB 4; Length 4599;  
Best Local Similarity 53.6%; Pred. No. 1.5e-25;  
Matches 45; Conservative 12; Mismatches 27; Indels 0; Gaps 0;  
  
QY 3 CSPKQACRQDITCISKWRCDCGERDCPDGSDPEAPICPQSKAQRCPQNEHNCGLGTELVCV 62  
Db 32 CDGFEFLCHDHVTCVQSQWLCDGDPDPCDDSDSLDTCPEVEIKPLNHIACLGKNCV 91  
  
QY 63 PMSRLNGVQDCMDGSDGPHCRE 86  
Db 92 HLSQLGNGVLDCPDGVDGPHCRE 115  
  
RESULT 8  
Q9J118 PRELIMINARY; PRT; 4599 AA.  
ID Q9J118  
AC Q9J118  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Low density lipoprotein receptor related protein LRP1b/LRP-DIT.  
GN LRP1b.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Yaklichkin S., Lisitsyn N.;  
RT "The structure of the mouse homologue of the human candidate tumor  
suppressor gene LRP1b/LRP-DIT."  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF270884; AAF81758.1; -  
DR HSP; Q07954; ICR8  
DR MGD; MGI-2151136; Lrp1b.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR00152; Asx hydroxyl\_S.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR006209; EGF\_like.  
DR InterPro; IPR002172; LDL\_receptor\_A.  
DR InterPro; IPR000033; LDL\_receptor\_rep.  
DR InterPro; IPR000169; SHprot\_acsite.  
DR Pfam; PF00008; EGF; 15.  
  
QY 3 CSPKQACRQDITCISKWRCDCGERDCPDGSDPEAPICPQSKAQRCPQNEHNCGLGTELVCV 62  
Db 32 CDGFEFLCHDHVTCVQSQWLCDGDPDPCDDSDSLDTCPEVEIKPLNHIACLGKNCV 91  
  
QY 63 PMSRLNGVQDCMDGSDGPHCRE 86  
Db 92 HLSQLGNGVLDCPDGVDGPHCRE 115  
  
RESULT 9  
Q46131 PRELIMINARY; PRT; 883 AA.  
ID Q46131  
AC Q46131  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Lipophorin receptor precursor.  
OS Locusta migratoria (Migratory locust).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;  
OC Acridoidea; Acrididae; Oedipodinae; Locusta.  
OX NCBI\_TaxID=7004;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Dantuma N.P., Potters M., Bogerd J., Vanderhorst D.J.;  
RT "An insect homolog of the vertebrate very low density lipoprotein  
receptor mediates endocytosis of lipophorins."  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ000010; CAA03855.1; -  
DR HSP; P01130; ILDR.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR InterPro; IPR000152; Asx hydroxyl\_S.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR006209; EGF\_like.  
DR InterPro; IPR002172; LDL\_receptor\_A.  
DR InterPro; IPR000033; LDL\_receptor\_rep.  
DR Pfam; PF00057; ldl\_recept\_a; 8.  
DR Pfam; PF00058; ldl\_recept\_b; 5.  
DR PRINTS; PR00261; LDLRECEPTOR.  
DR SMART; SM00179; EGF\_CA; 1.  
DR SMART; SM00192; LDLA; 8.  
DR SMART; SM00135; LY; 5.  
DR PROSITE; PS00010; ASX HYDROXYL; 2.  
DR PROSITE; PS01186; EGF\_2; 1.  
DR PROSITE; PS01187; EGF\_CA; 1.  
DR PROSITE; PS01209; LDLRA\_1; 6.  
DR PROSITE; PS00068; LDLRA\_2; 8.  
KW EGF-like domain; Receptor; Signal.  
FT SIGNAL 1 33 POTENTIAL.  
FT CHAIN 34 883 LIPOPHORIN RECEPTOR.

Pfam; PF00057; ldl\_recept\_a; 32.  
Pfam; PF00058; ldl\_recept\_b; 33.  
PRINTS; PR00261; LDLRECEPTOR.  
SMART; SM00179; EGF\_CA; 4.  
SMART; SM00192; LDLA; 32.  
SMART; SM00135; LY; 33.  
PROSITE; PS00010; ASX HYDROXYL; 4.  
PROSITE; PS00022; EGF\_1; 5.  
PROSITE; PS01186; EGF\_2; 9.  
PROSITE; PS01187; EGF\_CA; 3.  
PROSITE; PS01209; LDLRA\_1; 27.  
PROSITE; PS00068; LDLRA\_2; 32.  
PROSITE; PS00139; THIOLE PROTEASE\_CYS; 1.  
KW EGF-like domain; Lipoprotein; Receptor.  
SQ SEQUENCE 4599 AA; 513623 MW; 11462A3354FFB200 CRC64;  
  
Query Match 53.4%; Score 270; DB 11; Length 4599;  
Best Local Similarity 51.2%; Pred. No. 2.6e-24;  
Matches 43; Conservative 13; Mismatches 28; Indels 0; Gaps 0;  
  
QY 3 CSPKQACRQDITCISKWRCDCGERDCPDGSDPEAPICPQSKAQRCPQNEHNCGLGTELVCV 62  
Db 32 CDGFEFLCHDHVTCVQSQWLCDGDPDPCDDSDSLDTCPEVEIKPLNHIACLGSSACV 91  
  
QY 63 PMSRLNGVQDCMDGSDGPHCRE 86  
Db 92 HLSKLCNGVDCPDGFDGEGHCRE 115  
  
RESULT 9  
Q46131 PRELIMINARY; PRT; 883 AA.  
ID Q46131  
AC Q46131  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Lipophorin receptor precursor.  
OS Locusta migratoria (Migratory locust).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;  
OC Acridoidea; Acrididae; Oedipodinae; Locusta.  
OX NCBI\_TaxID=7004;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Dantuma N.P., Potters M., Bogerd J., Vanderhorst D.J.;  
RT "An insect homolog of the vertebrate very low density lipoprotein  
receptor mediates endocytosis of lipophorins."  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ000010; CAA03855.1; -  
DR HSP; P01130; ILDR.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR InterPro; IPR000152; Asx hydroxyl\_S.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR006209; EGF\_like.  
DR InterPro; IPR002172; LDL\_receptor\_A.  
DR InterPro; IPR000033; LDL\_receptor\_rep.  
DR Pfam; PF00057; ldl\_recept\_a; 8.  
DR Pfam; PF00058; ldl\_recept\_b; 5.  
DR PRINTS; PR00261; LDLRECEPTOR.  
DR SMART; SM00179; EGF\_CA; 1.  
DR SMART; SM00192; LDLA; 8.  
DR SMART; SM00135; LY; 5.  
DR PROSITE; PS00010; ASX HYDROXYL; 2.  
DR PROSITE; PS01186; EGF\_2; 1.  
DR PROSITE; PS01187; EGF\_CA; 1.  
DR PROSITE; PS01209; LDLRA\_1; 6.  
DR PROSITE; PS00068; LDLRA\_2; 8.  
KW EGF-like domain; Receptor; Signal.  
FT SIGNAL 1 33 POTENTIAL.  
FT CHAIN 34 883 LIPOPHORIN RECEPTOR.

```

SQ SEQUENCE 883 AA; 98332 MW; 86D0056ADB0A405B CRC64;
Query Match 41.3%; Score 209; DB 5; Length 883;
Best Local Similarity 46.2%; Pred. No. 2e-17;
Matches 36; Conservative 12; Mismatches 26; Indels 4; Gaps 2;

QY 3 CSPKQAFACDQITCISGKWRCDGRCPCDPSGDEAPETCPQSKAQRCQPNHNC-LGTEL 62
Db 255 CUPREFECLDRMTCHQSWCVDGRDPCDPSGDESVSRCHN---MTCRPDQFC-RNRICI 310
QY 63 PMSRLCNGVQDCWDGSD 80
Db 311 PGLHCSGHADCSGSD 328

RESULT 10
Q9VBN1 PRELIMINARY; PRT; 883 AA.
AC Q9VBN1
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG31094-PA.
GN C34861.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
CX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.J., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfankoch C., Baldwin D.,
RA Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostali D., Houson K.A., Howland T.J., Wei M.H., Ibegwan C.,
RA Jialali B., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,

```

```

RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Bazon J., An H., Baldwin D., Banon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwan C., Jallali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacle J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskaas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochownik S.E., Smith C.D.,
RA Tuzy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E., S.E.;
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003753; AAP56499.2; -
DR FlyBase; FB00039363; CG4861.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR000152; Asx hydroxyl_S.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR008210; IEGF.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000033; LDL_receptor_rep.
DR Pfam; PF00057; ldl_recept_a; 7.
DR Pfam; PF00058; ldl_recept_b; 5.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00181; EGF; 5.
DR SMART; SM00179; EGF_CA; 2.
DR SMART; SM00192; LDLra; 7.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS01209; LDLRA_1; 7.
DR PROSITE; PS00068; LDLRA_2; 7.
SQ SEQUENCE 883 AA; 98319 MW; 8245B1A3E9393F0 CRC64;

Query Match 41.0%; Score 207.5; DB 5; Length 883;
Best Local Similarity 50.0%; Pred. No. 3.1e-17;
Matches 40; Conservative 11; Mismatches 24; Indels 5; Gaps 3;

QY 2 TCSPKQAFACDQITCISGKWRCDGRCPCDPSGDEAPETCPQSKAQRCQPNHNC-LGTEL 60
Db 30 TCSSDDQFCNGG-NCIPKWRCDQESDCADGSDANELC---RARTCSPDEYACKSGEG 85
QY 61 CVPMSRLCNGVQDCWDGSD 80
Db 86 CVPLAWMCDQSKDCGSDGSD 105

RESULT 11
Q9V383 PRELIMINARY; PRT; 4699 AA.
ID Q9V383
AC Q9V383;

```

DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE CG8706 protein.  
GN CG33087 OR CG8706 OR CG33068.  
OS Drosophila melanogaster (Fruit fly)  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
ON NCBI\_TaxID=7227;  
RX STRAIN=Berkely;  
RP SEQUENCE FROM N.A.  
RC MEDLINE=2019606; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Ananidis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers J.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Arlil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherty J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Fodor C., Gabrielson A.E., Garg N.S., Gelbart W.M., Glassman K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E.C., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao O.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RL "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
[2]  
RP SEQUENCE FROM N.A.  
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Ananidis P.G., Brandon R.C., Rogers Y.,  
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
RA Ferrieria S., Frise E., Galle R.F., Garq N.S., George R.A.,  
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,  
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,  
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
RA Stapleton M., Strong R., Svirskaas R., Tector C., Tyler D.,  
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
RT "Sequencing of Drosophila melanogaster genomes.";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
[3]  
RP SEQUENCE FROM N.A.  
RA M.Sra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
RA Hradscky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,

Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,  
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,  
RA Seale S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,  
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
RT "Annotation of Drosophila melanogaster genome.";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
[4]  
RP SEQUENCE FROM N.A.  
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
[5]  
RP SEQUENCE FROM N.A.  
RA FlyBase;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE003839; AAF59114.3; -  
DR HSSP; P01130; 1AJJ.  
DR FlyBase; FBgn0053087; CG33087.  
DR GO; GO:0015020; C:membrane; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti.; IEA.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR003439; ABC\_transporter.  
DR InterPro; IPR00152; Asx\_hydroxyl\_s.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR002172; LDL\_receptor\_A.  
DR InterPro; IPR000033; LDL\_receptor\_reg.  
DR Pfam; PF00008; EGF; 12.  
DR Pfam; PF00057; ldl\_recept\_a; 27.  
DR Pfam; PF00058; ldl\_recept\_b; 27.  
DR PRINTS; PR00261; LDLRECEPTOR.  
DR SMART; SM00181; EGF; 39.  
DR SMART; SM00179; EGF\_CA; 13.  
DR SMART; SM00192; LDLA; 31.  
DR SMART; SM00135; LY; 31.  
DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 1.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 4.  
DR PROSITE; PS00022; EGF\_1; 8.  
DR PROSITE; PS01186; EGF\_2; 9.  
DR PROSITE; PS01187; EGF\_CA; 3.  
DR PROSITE; PS01209; LDLRA\_1; 22.  
DR PROSITE; PS00068; LDLRA\_2; 29.  
DR EGF-like domain.  
KW EGF-like domain.  
SQ SEQUENCE 4699 AA; 524916 MW; 166D46B38C74A6BF CRC64;  
Query Match 40.6%; Score 205.5; DB 5; Length 4699;  
Best Local Similarity 45.8%; Pred.No.3e-16;  
Matches 38; Conservative 9; Mismatches 27; Indels 9; Gaps 3;  
QY 2 TCSPKQAFACRDQITCISKWECDCGERDCPDGSDPEAIPCQSKAQRCQPNHNCGLGTCLC 61  
Db 2756 SCDFQGFACSG-NCIAGSWECDCGERDCPDGSDPEI-----NCRTECHNQFAC--DKTC 2806  
QY 62 VPMRLCNGVQDCMDGSDGPHC 84  
Db 2807 IPASWQCCKGSDCEDGSDGPHC 2829  
RESULT 12  
QYW57  
ID Q7YW57 PRELIMINARY; PRT; 891 AA.  
AC Q7YW57;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Lipophorin receptor.  
OS Aedes aegypti (Yellowfever mosquito).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes.  
ON NCBI\_TaxID=7159;

```

RN      SEQUENCE FROM N.A.
RP      Seq S.-J., Jun H.-M., Sun J., Sappington T.W., Raikhel A.S.;
RT      "Tissue- and stage-specific expression of two lipophorin receptor
RT      variants with seven and eight ligand-binding repeats in the adult
RT      mosquito.";
RL      Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AY348869; AAQ16410.1; -.
KW      Receptor.
SQ      SEQUENCE. 891 AA; 99303 MW; 82D411CD45EB271A CRC64;

Query Match      40.4%; Score 204.5; DB 5; Length 891;
Best Local Similarity 44.9%; Pred. No. 7.3e-17;
Matches 35; Conservative 11; Mismatches 29; Indels 3; Gaps 1;

QY      3 CSPKQACRDQITCISKGRCDGERCPDGSDEAPEICPSKAQRCPQNEHNCGLGTCLV 62
DB      205 CLSLEYQCSDRITCIHKSWICDGEKDCPQGDDEMPPICT---QNVTCRPDQFQCKDKTCT 261

QY      63 PMSRLCNGVQDCMDGSD 80
DB      262 NGHFHCNGKPECSDGSD 279

RESULT 13
Q963T3      PRELIMINARY; PRT; 1156 AA.
AC      Q963T3;
DT      01-DEC-2001 (T-EMBLrel. 19, Created)
DT      01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT      01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE      Lipophorin receptor.
OS      Aedes aegypti (Yellow fever mosquito).
OC      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC      Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes.
OX      NCBI_TaxID=7159;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=21272543; PubMed=11378410;
RA      Cheon H., Seo S., Sun J., Sappington T.W., Raikhel A.S.;
RT      "Molecular characterization of the VLDL receptor homolog mediating
RT      binding of lipophorin in oocyte of the mosquito Aedes aegypti.";
RL      Insect Biochem. Mol. Biol. 31:753-760(2001).
DR      EMBL: AF355595; AAK72954.1; -.
DR      GO: GO:0016020; C:membrane; IEA.
DR      GO: GO:0005509; F:calcium ion binding; IEA.
DR      GO: GO:0004872; F:receptor activity; IEA.
DR      InterPro: IPR000152; Asx_hydroxyl_S.
DR      InterPro: IPR001881; EGF_Ca.
DR      InterPro: IPR006209; EGF-like.
DR      InterPro: IPR002172; LDL_receptor_A.
DR      InterPro: IPR000033; Ldl_receptor_rep.
DR      Pfam: PF00057; ldl_recept_a; 8.
DR      PRINTS: PR00058; ldl_recept_b; 5.
DR      SMART: SM00179; EGF_Ca; 1.
DR      SMART: SM00192; LDLRA_2; 8.
DR      SMART: SM00135; LY; 5.
DR      PROSITE: PS00010; ASX_HYDROXYL; 1.
DR      PROSITE: PS01186; EGF_2; 1.
DR      PROSITE: PS01209; LDLRA_1; 7.
DR      PROSITE: PS50068; LDLRA_2; 8.
KW      Receptor.
SQ      SEQUENCE. 1156 AA; 128931 MW; 050D55E6C860E30F CRC64;

Query Match      40.4%; Score 204.5; DB 5; Length 1156;
Best Local Similarity 44.9%; Pred. No. 9.6e-17;
Matches 35; Conservative 11; Mismatches 29; Indels 3; Gaps 1;

QY      3 CSPKQACRDQITCISKGRCDGERCPDGSDEAPEICPSKAQRCPQNEHNCGLGTCLV 62
DB      299 CLSLEYQCSDRITCIHKSWICDGEKDCPQGDDEMPPICT---QNVTCRPDQFQCKDKTCT 355

QY      63 PMSRLCNGVQDCMDGSD 80
DB      356 NGHFHCNGKPECSDGSD 373

RESULT 14
Q8WY29      PRELIMINARY; PRT; 1537 AA.
AC      Q8WY29;
DT      01-MAR-2002 (T-EMBLrel. 20, Created)
DT      01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT      01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE      Candidate tumor suppressor protein (Fragment).
GN      LRP1B.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=20487554; PubMed=11031110;
RA      Liu C.-X., Musco S., Lisitsyna N.M., Yaklichkin S.Y., Lisitsyn N.A.;
RT      "Genomic organization of a new candidate tumor suppressor gene,
RT      LRP1B";
RL      Genomics 69:271-274(2000).
DR      EMBL: AF83408; AAL38109.1; -.
DR      EMBL: AF83376; AAL38109.1; JOINED.
DR      EMBL: AF83377; AAL38109.1; JOINED.
DR      EMBL: AF83378; AAL38109.1; JOINED.
DR      EMBL: AF83379; AAL38109.1; JOINED.
DR      EMBL: AF83380; AAL38109.1; JOINED.
DR      EMBL: AF83381; AAL38109.1; JOINED.
DR      EMBL: AF83382; AAL38109.1; JOINED.
DR      EMBL: AF83383; AAL38109.1; JOINED.
DR      EMBL: AF83384; AAL38109.1; JOINED.
DR      EMBL: AF83385; AAL38109.1; JOINED.
DR      EMBL: AF83386; AAL38109.1; JOINED.
DR      EMBL: AF83387; AAL38109.1; JOINED.
DR      EMBL: AF83388; AAL38109.1; JOINED.
DR      EMBL: AF83389; AAL38109.1; JOINED.
DR      EMBL: AF83390; AAL38109.1; JOINED.
DR      EMBL: AF83391; AAL38109.1; JOINED.
DR      EMBL: AF83392; AAL38109.1; JOINED.
DR      EMBL: AF83393; AAL38109.1; JOINED.
DR      EMBL: AF83394; AAL38109.1; JOINED.
DR      EMBL: AF83395; AAL38109.1; JOINED.
DR      EMBL: AF83396; AAL38109.1; JOINED.
DR      EMBL: AF83397; AAL38109.1; JOINED.
DR      EMBL: AF83398; AAL38109.1; JOINED.
DR      EMBL: AF83399; AAL38109.1; JOINED.
DR      EMBL: AF83400; AAL38109.1; JOINED.
DR      EMBL: AF83401; AAL38109.1; JOINED.
DR      EMBL: AF83402; AAL38109.1; JOINED.
DR      EMBL: AF83403; AAL38109.1; JOINED.
DR      EMBL: AF83404; AAL38109.1; JOINED.
DR      EMBL: AF83405; AAL38109.1; JOINED.
DR      EMBL: AF83406; AAL38109.1; JOINED.
DR      EMBL: AF83407; AAL38109.1; JOINED.
DR      GO: GO:0016020; C:membrane; IEA.
DR      GO: GO:0005509; F:calcium ion binding; IEA.
DR      InterPro: IPR000152; Asx_hydroxyl_S.
DR      InterPro: IPR001881; EGF_Ca.
DR      InterPro: IPR006209; EGF-like.
DR      InterPro: IPR000033; Ldl_receptor_rep.
DR      Pfam: PF00057; ldl_recept_a; 18.
DR      Pfam: PF00058; ldl_recept_b; 9.
DR      PRINTS: PR00261; LDLRECEPTOR.
DR      SMART: SM00179; EGF_Ca; 3.
DR      SMART: SM00152; LDLRA; 18.
DR      SMART: SM00135; LY; 8.
DR      PROSITE: PS00010; ASX_HYDROXYL; 3.
DR      PROSITE: PS01186; EGF_2; 2.
```

DR PROSITE; PS01187; EGF\_CA; 2.  
DR PROSITE; PS01209; LDLRA\_1; 14.  
DR PROSITE; PS00068; LDLRA\_2; 18.  
KW EGF-like domain.  
FT NON\_TER 1  
FT NON\_TER 1537 1537  
SQ SEQUENCE 1537 AA; 172525 MW; 6F38AAFC818F135C CRC64;

Query Match  
Best Local Similarity 51.9%; Score 203; DB 4; Length 1537;  
Matches 41; Conservative 7; Mismatches 27; Indels 4; Gaps 2;  
QY 2 TCSPKQFACRQDQITCISKWRCGGERDCPDGSDPEAIPCQSKAQCQPNHNCILGTEL 61  
DB 43 SCWNQFACSAQ-KCISKWICGDDCGDLDESISIC---CAITCAADMFSCQGSRA 98  
QY 62 VPMRLCNGVQDCMDGSDSDE 80  
DB 99 VPRHLCGGERDCPDGSDSDE 117

## RESULT 15

Q81GR9 Q81GR9 PRELIMINARY; PRT; 739 AA.  
AC Q81GR9;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE R33884P.  
GN CG4861.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Y;  
RA Stapleton M., Brokston P., Hong L., Agbavani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,  
RA Miranda A., Mungall C.J., Nuroo J., Pacleb J., Paragas V., Park S.,  
RA Patel S., Prounanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
RA Ceiniker S.;  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BT001632; AAN71387.1; -.  
DR FlyBase; FBgn039363; CG4861.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR InterPro; IPR000152; Asx\_Hydroxyl\_S.  
DR InterPro; IPR001881; EGF\_CA.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR006210; HGF.  
DR InterPro; IPR002172; LDL\_receptor\_A.  
DR InterPro; IPR000033; Ldl\_receptor\_rep.  
DR Pfam; PF00009; EGF; 1.  
DR Pfam; PF00057; ldl\_recept\_a; 5.  
DR Pfam; PF00058; ldl\_recept\_b; 5.  
DR PRINTS; PR00261; LDLRECEPTOR.  
DR SMART; SM00181; EGF; 4.  
DR SMART; SM00179; EGF\_CA; 2.  
DR SMART; SM00132; LDL; 5.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 1.  
DR PROSITE; PS01186; EGF\_2; 1.  
DR PROSITE; PS01187; EGF\_CA; 1.  
DR PROSITE; PS01209; LDLRA\_1; 5.  
DR PROSITE; PS00068; LDLRA\_2; 5.  
SQ SEQUENCE 739 AA; 82203 MW; 53162F0A9EF82823 CRC64;

Query Match  
Best Local Similarity 44.9%; Score 202; DB 5; Length 739;  
Matches 35; Conservative 11; Mismatches 28; Indels 4; Gaps 2;

QY 3 CSPKQFACRQDQITCISKWRCGGERDCPDGSDPEAIPCQSKAQCQPNHNCILGTEL 62  
DB 105 CLSHEYQCKDRITCLHHSWLCGDRDCPDGDEHTANC---KNVTCRADQFQC-GDRSCI 160  
QY 63 PMSRLCNGVQDCMDGSDSDE 80  
DB 161 PGHLTCNGDKDCADGSDSDE 178  
Search completed: September 17, 2004, 11:17:54  
Job time : 119 secs

This Page Blank (uspto)